COVERING MOVES AND KIRBY CALCULUS

I. Bobtcheva

Dipartimento di Scienze Matematiche
Università Politecnica delle Marche – Italia

`bobtchev@dipmat.unian.it`

R. Piergallini

Dipartimento di Matematica e Informatica
Università di Camerino – Italia

`riccardo.piergallini@unicam.it`

Abstract

We show that simple coverings of B^4 branched over ribbon surfaces up to certain local ribbon moves bijectively represent orientable 4-dimensional 2-handlebodies up to handle sliding and addition/deletion of cancelling handles. As a consequence, we obtain an equivalence theorem for simple coverings of S^3 branched over links, in terms of local moves. This result generalizes to coverings of any degree results by the second author and Apostolakis, concerning respectively the case of degree 3 and 4. We also provide an extension of our equivalence theorem to possibly non-simple coverings of S^3 branched over embedded graphs.

This work represents the first part of our study of 4-dimensional 2-handlebodies. In the second part [9], we factor such bijective correspondence between simple coverings of B^4 branched over ribbon surfaces and orientable 4-dimensional 2-handlebodies through a map onto the closed morphisms in a universal braided category freely generated by a Hopf algebra object.

Keywords: 3-manifold, 4-manifold, branched covering, branching link, branching graph, branching ribbon surface, covering move, ribbon move, Kirby calculus.

AMS Classification: 57M12, 57M25, 57N10, 57N13, 57Q45.

Introduction

In the early 70's Hilden [17, 18], Hirsch [19] and Montesinos [28, 29] independently proved that every closed connected oriented 3-manifold can be represented as a 3-fold simple covering of S^3 branched over a link. Successively, Montesinos [31] obtained an analogous representation of any connected oriented 4-manifold admitting a finite handlebody decomposition with handles of indices ≤ 2 as a simple 3-fold covering of B^4 branched over a possibly non-orientable ribbon surface. Actually, the branching surface can always be made orientable as we remark at the end of Section 2 (cf. [26, 39] for other constructions giving directly orientable ribbon surfaces).

The problem of finding moves relating any two such covering representations of the same manifold was first considered by Montesinos. For the 3-dimensional case, in [34] he proposed the two local moves M_1 and M_2 of Figure 1, where i, j, k and l are all distinct, in terms of branching links and monodromy. Here, as well as in all the following pictures of moves, we draw only the part of the labelled branching set inside the relevant cell, assuming it to be fixed outside this cell.
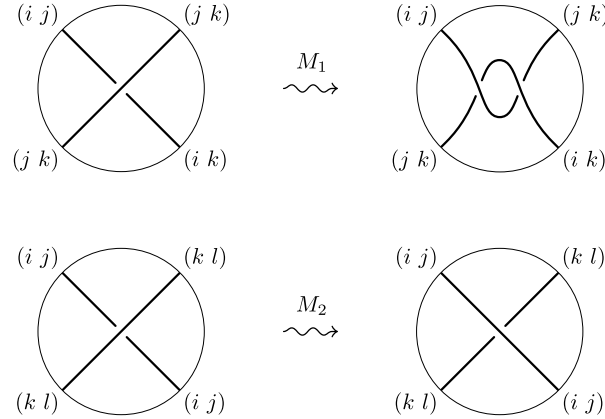

FIGURE 1.

It is worth observing that the inverse move M_1^{-1} can be realized, up to labelled isotopy, by a composition of two moves M_1. We leave this easy exercise to the reader, referring to Figure 11 of [38] for the solution. On the other hand the inverse move M_2^{-1} coincides up to isotopy with the move M_2, becoming distinct from it only after an orientation is fixed on the branching link.

A complete set of moves for 3-fold simple coverings of S^3 branched over a link was given in [37] by the second author. Such moves are non-local, but in [38] the local moves M_1 and M_2 are shown to suffice after stabilization with a fourth trivial sheet. In [38] the question was also posed, whether these local moves together with stabilization suffice for covering representations of arbitrary degree. Recently, Apostolakis [4] answered this question positively for coverings of degree 4.
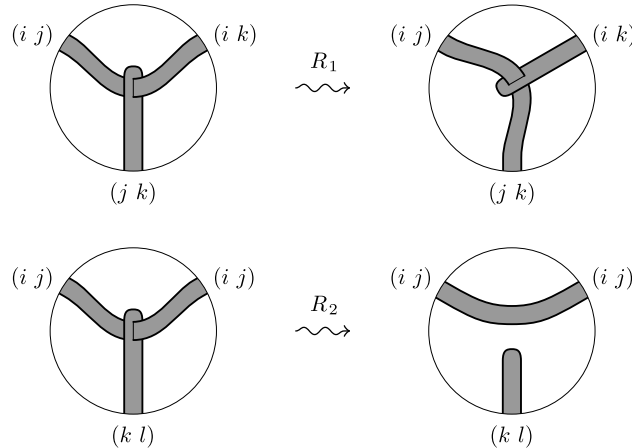

FIGURE 2.

In this paper, we derive the solution of the moves problem for arbitrary degree simple coverings of S^3 branched over links (cf. Theorem 3), from an equivalence theorem for simple coverings of B^4 branched over ribbon surfaces, that relates the local ribbon moves R_1 and R_2 of Figure 2, where i, j, k and l are all distinct, with the 4-dimensional Kirby calculus (cf. Theorem 1). In particular, our result does not depend on the partial ones of [37], [38] and [4].

Analogously to the Montesinos moves, also these ribbon moves generate their inverses up to labelled isotopy (cf. Proposition 2.5). This is obvious for the move R_1, if we think of it as rotation of $120°$ (followed by relabelling), being $R_1^{-1} = R_1^2$. We leave to the reader to verify that R_2^{-1} coincide with R_2 up to labelled isotopy (they become distinct once the branching ribbon surface is oriented).

Given a connected simple covering $p : M \to B^4$ branched over a ribbon surface $F \subset B^4$, we have that any 2-dimensional 1-handlebody structure on F induces a 4-dimensional 2-handlebody structure on M (see Section 1 for the definition of m-dimensional n-handlebody). In fact, the simple covering of B^4 branched over the disjoint union of trivial disks F_0, representing the 0-handles of F, can be easily seen to be a 4-dimensional 1-handlebody M_1. Moreover, following [31] (cf. also [20]), any 1-handle of F attached to F_0 corresponds to a 2-handle of M attached to M_1.

In Section 2 we show that handle sliding and handle cancellation in F give raise to analogous modifications in M. Therefore, the 2-handlebody structure of M turns out to be uniquely determined by the labelled ribbon surface F up to 2-equivalence, that is up to handle sliding and addition/deletion of cancelling pairs of handles of indices ≤ 2 (cf. Section 1). In other words, any simple covering of B^4 branched over a ribbon surface represents a well defined 2-equivalence class of 4-dimensional 2-handlebodies.

The main result of Montesinos [31] is that any connected oriented 4-dimensional 2-handlebody M has a 3-fold branched covering representation as above. The corresponding labelled ribbon surface F, with the right 2-dimensional 1-handlebody structure, is obtained from a Kirby diagram of M, after it has been suitably symmetrized with respect to a standard 3-fold simple covering representation of M_1.

In Section 3 (see also Remark 4.4) we give a different construction of the labelled ribbon surface F, similar to that one of labelled links given in [32] for 3-manifolds (cf. Remark 3.4). Our construction is simpler and more effective than the Montesinos one, is canonical up to ribbon moves and better preserves the structure of the starting Kirby diagram, allowing us to interpret the Kirby calculus in terms of ribbon moves.

At this point, we are ready to state our first theorem. In substance, it asserts that simple coverings of B^4 branched over ribbon surfaces up certain local isotopy moves, stabilization and ribbon moves R_1 and R_2 bijectively represent 4-dimensional 2-handlebodies up to 2-equivalence. For the sake of simplicity, we consider only the connected case. Nevertheless, as we remark at the end of this introduction, the statement essentially holds in the general case too, provided the lower bound for the stabilization degree is replaced by the appropriate one (cf. Proposition 4.5).

THEOREM 1. *Two connected simple coverings of B^4 branched over ribbon surfaces represent 2-equivalent 4-dimensional 2-handlebodies if and only if after stabilization to the same degree ≥ 4 their labelled branching surfaces can be related by labelled 1-isotopy and a finite sequence of moves R_1 and R_2.*

The definition of 1-isotopy is given in Section 1. Here, we limit ourselves to say that it is essentially generated by the local isotopy moves shown in Figure 3 (cf. Proposition 1.3). We do not know whether 1-isotopy coincides with isotopy of ribbon surfaces (see discussion in Sections 1 and 5). Anyway, we have 1-isotopy instead of isotopy in the statement of Theorem 1, due to Lemma 2.3. The proof of the theorem is achieved in Section 4, as a consequence of the above mentioned covering representation of Kirby calculus. Other main ingredients are Propositions 4.2 and 4.3.
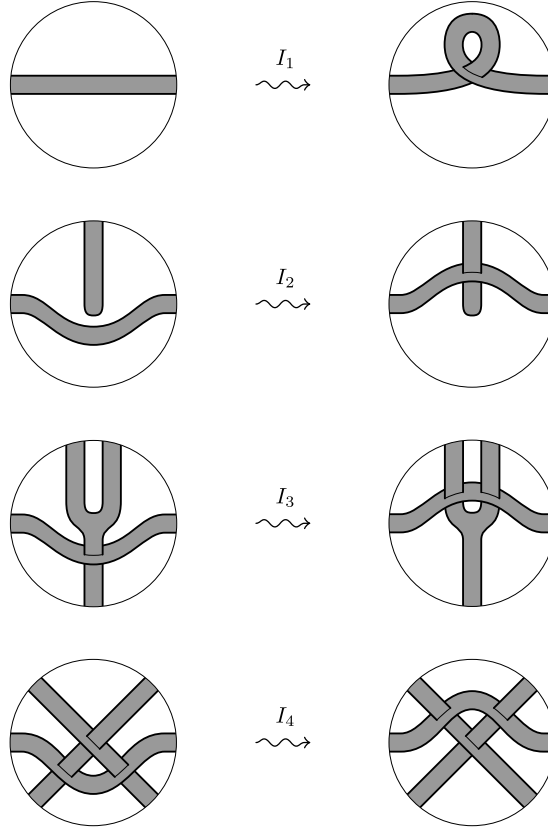

FIGURE 3.

Now, in order to deal with 3-manifolds, we need to introduce the further moves depicted in Figures 4 and 5. As we see in Section 4, these moves allow us to realize respectively positive/negative blow up and handle trading. In particular, they are not covering moves in the sense defined in Section 1, since they change the covering 4-manifold. On the other hand, they do not change the restriction of the covering over S^3, leaving the boundary of the branching surface fixed up to isotopy.

The next theorem, whose proof is given in Section 4, tells us that these last moves together with their inverses and the previous ribbon moves suffice to completely represent the Kirby calculus for 3-manifolds. Notice that here, differently from the statement of Theorem 1, labelled isotopy can be equivalently used instead of labelled 1-isotopy, since it preserves the covering manifold up to diffeomorphism.

THEOREM 2. *Two connected simple coverings of B^4 branched over ribbon surfaces represent 4-manifolds with diffeomorphic oriented boundaries if and only if*

after stabilization to the same degree ≥ 4 their labelled branching surfaces can be related by labelled isotopy and a finite sequence of moves R_1, R_2, $P_{\pm}^{\pm 1}$ and $T^{\pm 1}$.
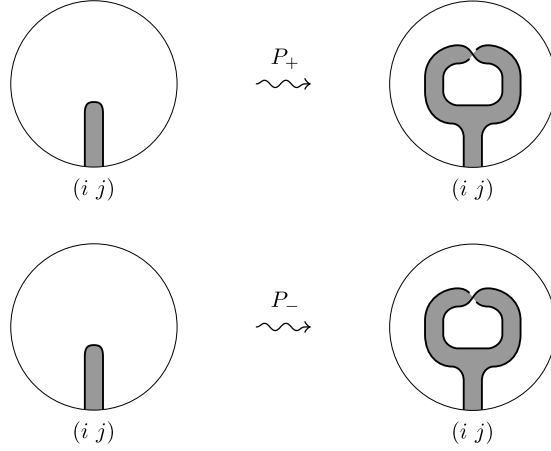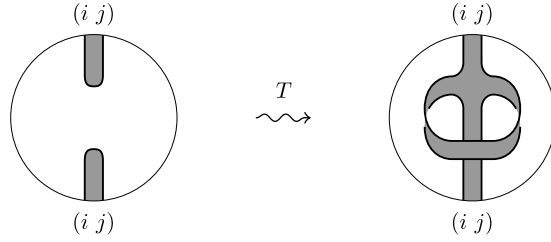

FIGURE 4.

FIGURE 5.

By focusing on the boundary, we observe that the restrictions of the ribbon moves R_1 and R_2 to S^3 can be realized respectively by Montesinos moves M_1 and M_2. This is shown in Figure 6 for move R_1, while it is trivial for move R_2. In both cases we can apply two Montesinos moves inverse to each other (with respect to any local orientation of the link as boundary of the surface, for move M_2).
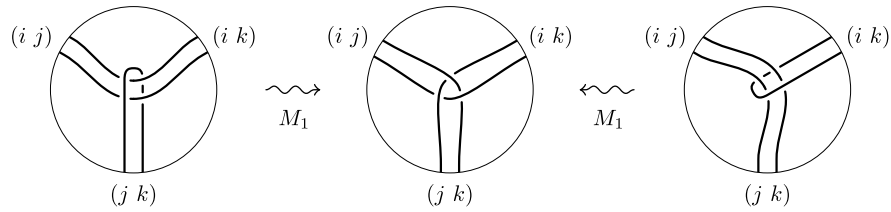

FIGURE 6.

This observation allows us to derive from Theorem 2 the following theorem for simple covering of S^3 branched over links.

THEOREM 3. *Two connected simple coverings of S^3 branched over links represent diffeomorphic oriented 3-manifolds if and only if after stabilization to the same degree ≥ 4 their labelled branching links can be related by labelled isotopy and a finite sequence of moves M_1 and M_2.*

We prove Theorem 3 in Section 4, as a consequence of Proposition 4.7. This says that any labelled link representing a simple branched covering of S^3 can be transformed through Montesinos moves into the boundary of a labelled ribbon surface representing a simple branched covering of B^4.

Finally, we want to extend Theorem 3 to arbitrary branched coverings of S^3. To do that, we introduce the moves S_1 and S_2 depicted in Figure 7. Here, the branching set is allowed to be singular and the monodromy is not necessarily simple. In fact, σ_1 and σ_2 are any permutations, coherent in the sense defined Section 1, and $\sigma = \sigma_1\sigma_2$. This is the reason why we need to specify orientations for the arcs or equivalently positive meridians to which refer the monodromies.
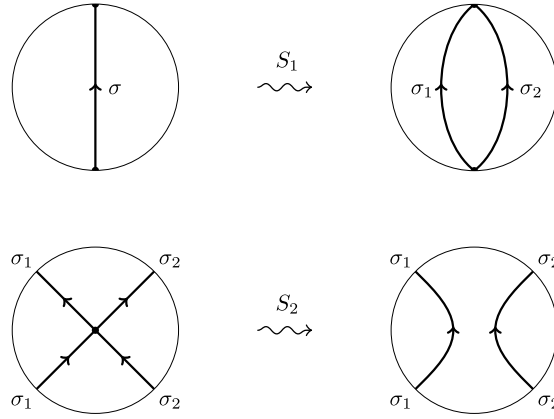

FIGURE 7.

Our last theorem is the wanted extension of Theorem 3. Its proof, given in Section 4, is based on the fact that moves S_1 and S_2 suffice to make simple any branched covering of S^3 and to remove all the singularities from its branching set.

THEOREM 4. *Two connected coverings of S^3 branched over a graph represent diffeomorphic oriented 3-manifolds if and only if after stabilization to the same degree ≥ 4 their branching graphs can be related by labelled isotopy and a finite sequence of moves M_1, M_2, $S_1^{\pm 1}$, $S_2^{\pm 1}$.*

We notice that all the above theorems could be easily reformulated to deal with non connected branched coverings too. Since everything can be done componentwise, possibly after labelling conjugation, it obviously suffice to stabilize the coverings to have the same number of sheets ≥ 4 for corresponding components. Moreover, as will be clear at the end of Section 4, the total degree can be lowered to $3c + 1$, where c is the maximum number of components of the two coverings, if we allow stabilization/destabilization at intermediate stages (cf. Proposition 4.5).

In conclusion, it is also worth remarking that our results, beyond establishing a strong relation between branched covering presentations and Kirby diagrams of 3- and 4-manifolds, also provide an effective way to pass from one to the other. We discuss this aspect in Section 5.

1. Preliminaries

Before going into details, we fix some general notations and conventions about handlebodies, that will be used in various contexts in the following. We refer to [14] or [23] for all the definitions and basic results not explicitly mentioned here.

We recall that an *i-handle* of dimension m is a copy H^i of $B^i \times B^{m-i}$ attached to the boundary of an m-manifold M by an embedding $\varphi : S^{i-1} \times B^{m-i} \to \mathrm{Bd}\, M$. The two balls $B^i \times \{0\}$ and $\{0\} \times B^{m-i}$ in $M' = M \cup_\varphi H^i$ are called respectively the *core* and the *cocore* of H^i, while their boundaries $S^{i-1} \times \{0\}$ and $\{0\} \times S^{m-i-1}$ are called the *attaching sphere* and the *belt sphere* of H^i. Inside H^i, *longitudinal* means parallel to the core and *transversal* means parallel to the cocore. Up to isotopy, the attaching map φ is completely determined by the attaching sphere together with its *framing* in $\mathrm{Bd}\, M$, given by $S^{i-1} \times \{*\}$ for any $* \in B^{m-i} - \{0\}$.

Then, an *n-handlebody* of dimension m is defined by induction on n to be obtained by simultaneously smoothly attaching a finite number of n-handles to an $(n-1)$-handlebody of the same dimension m, starting with a disjoint union of 0-handles for $n = 0$.

By a well known result of Cerf [10] (cf. [14] or [23]), two handlebodies of the same dimension are diffeomorphic (forgetting their handle structure), if and only if they can be related by a finite sequence of the following modifications: 1) isotoping the attaching map of i-handles; 2) adding/deleting a pair of *cancelling handles*, that is a i-handle H^i and a $(i+1)$-handle H^{i+1}, such that the attaching sphere of H^{i+1} intersects the belt sphere of H^i transversally in a single point; 3) *handle sliding* of one i-handle H^i_1 over another one H^i_2, that means pushing the attaching sphere of H^i_1 through the belt sphere of H^i_2.

We call *k-deformation* any finite sequence of the above modifications such that at each stage we have an n-handlebody with $n \leq k$, that is we start from a n-handlebody with $n \leq k$ and never add any cancelling i-handle with $i > k$. Furthermore, we call *k-equivalent* two handlebodies related by a k-deformation.

In particular, any compact surface with non-empty boundary has a 1-handlebody structure and any two such structures are easily seen to be 1-equivalent (cf. proof of Proposition 1.2).

The other relevant case for our work is that one of orientable 4-manifolds (with non-empty boundary) admitting a 4-dimensional 2-handlebody structure. Any two such structures are 3-equivalent, but whether they are 2-equivalent is a much more subtle open question, which is expected to have negative answer (cf. Section I.6 of [23] and Section 5.1 of [14]). This question seems to be strongly related to the problem of finding isotopy moves for ribbon surfaces in B^4. In fact, as we will see, 4-dimensional 2/3-deformations correspond by means of branched coverings to regularly embedded 2-dimensional 1/2-deformations of branching surfaces in B^4 (cf. Proposition 2.2 and the discussion in Section 5).

Links

As usual, we represent a link $L \subset R^3 \subset R^3 \cup \infty \cong S^3$ by a planar *diagram* $D \subset R^2$, consisting of the orthogonal projection of L into R^2, that can be assumed self-transversal after a suitable horizontal (height preserving) isotopy of L, with a *crossing state* for each double point, telling which arc passes over the other one.

Such a diagram D uniquely determines L up to vertical isotopy. On the other hand, link isotopy can be represented in terms of diagrams by crossing preserving isotopy in R^2 and Reidemeister moves.

A link L is called *trivial* if it bounds a disjoint union of disks in R^3. It is well known that any link diagram D can be transformed into a diagram D' of a trivial link by suitable crossing changes, that is by inverting the state of some of its crossings. We say that D' is a *trivial state* of D. Actually, any link diagram D has many trivial states, but it is not clear at all how they are related to each other. For this reason, we are lead to introduce the more restrictive notions of vertically trivial link and vertically trivial state of a link diagram.

We say that a link L is *vertically trivial* if it meets any horizontal plane (parallel to R^2) in at most two points belonging to the same component. In this case, the height function separates the components of L (that is the height intervals of different components are disjoint), so that we can vertically order the components of L according to their height. Moreover, each component can be split into two arcs on which the height function is monotone, assuming the only unique minimum and maximum values at the common endpoints. Then, all the (possibly degenerate) horizontal segments spanned by L in R^3 form a disjoint union of disks bounded by L. This proves that L is a trivial link.

By a *vertically trivial state* of a link diagram D we mean any trivial state of D which is the diagram of a vertically trivial link. A vertically trivial state D' of D can be constructed by the usual naive unlinking procedure: 1) number the components of the link L represented by D and fix on each component an orientation and a starting point away from crossings; 2) order the points of L lexicographically according to the numbering of the components and then to the starting point and the orientation of each component; 3) resolve each double point of D into a crossings of D' by letting the arc which comes first in the order pass under the other one. The link L' represented by D' can be clearly assumed to be vertically trivial, considering on it a height function which preserves the order induced by the vertical bijection with L except for a small arc at the end of each component. Figure 8 (a) shows how the height function of a component looks like with respect to a parametrization having the starting point and the orientation fixed above. Keeping the parametrization fixed but changing the starting point or the orientation we get different height functions as in Figures 8 (b) and (c) respectively.

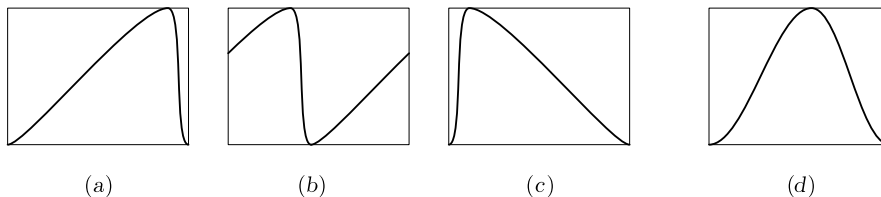(a) (b) (c) (d)

FIGURE 8.

Notice that the above unlinking procedure gives us only very special vertically trivial states. While it is clear how to pass from (a) to (b), by moving the starting point along the component, going from (a) to (c) turns out to be quite mysterious without considering generic vertically trivial states. The height function of a

component for such a state, with respect to a parametrization starting from the unique minimum point, looks like in Figure 8 (e), that is apparently an intermediate state between (a) and (c). The following proposition settles the problem of relating different vertically trivial states of the same link diagram.

PROPOSITION 1.1. *Any two vertically trivial states D' and D'' of a link diagram D are related by a sequence $D_0, D_1, \ldots, D_n$ of vertically trivial states of D, such that $D_0 = D'$, $D_n = D''$ and, for each $i = 1, \ldots, n$, D_i is obtained from D_{i-1} by changing a single self-crossing of one component or by changing all the crossings between two vertically adjacent components.*

Proof. Since the effect of changing all the crossings between two vertically adjacent components is the transposition of these components in the vertical order, by iterating this kind of modification we can permute as we want the vertical order of all the components. Hence, we only need to address the case of a knot diagram.

Given a knot diagram $D \subset R^2$ with double points $x_1, \ldots, x_n \in R^2$, we consider a parametrization $h : S^1 \to D$ and denote by $t_i', t_i'' \in S^1$ the two values of the parameter such that $h(t_i') = h(t_i'') = x_i$, for any $i = 1, \ldots, n$.

For any smooth knot $K \subset R^3$ which projects to a vertically trivial state of D, let $h_K : S^1 \to R^3$ be the parametrization of K obtained by lifting h and $f_K : S^1 \to R$ be the composition of h_K with the height function. Then, f_K is a smooth function with the following properties: 1) f_K has only one minimum and one maximum; 2) $f_K(t_i') \neq f_K(t_i'')$, for any $i = 1, \ldots, n$. In this way, the space of all smooth knots which project to vertically trivial states of D can be identified with the space of all smooth functions $f : S^1 \to R$ satisfying properties 1 and 2.

Now, the space $\mathcal{S}$ of all smooth functions $f : S^1 \to R$ satisfying property 1 is clearly pathwise connected, while the complement $\mathcal{C} \subset \mathcal{S}$ of property 2 is a closed codimension 1 stratified subspace. Therefore, if K' and K'' are knots projecting to the vertically trivial states D' and D'', then we can join $f_{K'}$ and $f_{K''}$ by a path in $\mathcal{S}$ transversal with respect to $\mathcal{C}$. This, path gives rise to a finite sequence of self-crossing changes as in the statement, one for each transversal intersection with $\mathcal{C}$. $\square$

We remark that the singular link between two consecutive vertically trivial states, obtained from each other by a single self-crossing change, is trivial. Namely, the unique singular component spans a 1-point union of two disks, disjoint from all the other components. This fact, which will play a crucial role in the proof of Proposition 3.1, follows from [42] but can also be easily proved directly by inspection.

Ribbon surfaces

A smooth compact surface $F \subset B^4$ with $\operatorname{Bd} F \subset S^3$ is called a *ribbon surface* if the Euclidean norm restricts to a Morse function on F with no local maxima in $\operatorname{Int} F$. Assuming $F \subset R_-^4 \subset R_-^4 \cup \{\infty\} \cong B^4$, this property is topologically equivalent to the fact that the fourth Cartesian coordinate restricts to a Morse height function on F with no local maxima in $\operatorname{Int} F$. Such a surface $F \subset R_-^4$ can be horizontally (preserving the height function) isotoped to make its orthogonal projection into R^3 a self-transversal immersed surface, whose double points form disjoint arcs as in Figure 9.
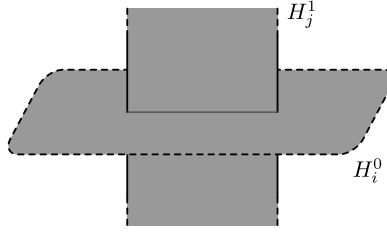
FIGURE 9.

We will refer to such a projection as a 3-dimensional *diagram* of F. Actually, any immersed compact surface $F \subset R^3$ with no closed components and all self-intersections of which are as above, is the diagram of a ribbon surface uniquely determined up to vertical isotopy. This can be obtained by pushing $\text{Int}\, F$ down inside $\text{Int}\, R^4_-$ in such a way that all self-intersections disappear.

In the following, *ribbon surfaces will be always represented by diagrams and considered up to vertical isotopy*. Moreover, *we will use the same notations for a ribbon surface and for its diagram in R^3*, disregarding the projection. By the above observation there will be no danger of confusion, provided that the ambient space will be clear.

Since a ribbon surface F has no closed components, it admits a handlebody decomposition $F = H_1^0 \cup \ldots \cup H_m^0 \cup H_1^1 \cup \ldots \cup H_n^1$ with only 0- and 1-handles. Such a 1-handlebody decomposition is called *adapted*, if each ribbon self-intersection involves an arc contained in the interior of a 0-handle and a proper transversal arc inside a 1-handle. (cf. [41]).

By an *embedded 2-dimensional 1-handlebody* we mean a ribbon surface endowed with an adapted 1-handlebody decomposition as above. Looking at the diagram, we have that the H_i^0's are disjoint non-singular disks, while the H_j^1's are non-singular bands attached to the H_i^0's and possibly passing across them as shown in Figure 9. Moreover, we can think of F as a smooth perturbation of the boundary of $((H_1^0 \cup \ldots \cup H_m^0) \times [0, -1]) \cup ((H_1^1 \cup \ldots \cup H_n^1) \times [0, -1/2])$, in such a way that the handlebody decomposition turns out to be induced by the height function.

We say that two embedded 2-dimensional 1-handlebodies are equivalent up to *embedded 1-deformation*, or briefly that they are *1-equivalent*, if they are related by a finite sequence of the following modifications:

(a) *adapted isotopy*, that is isotopy of 1-handlebodies in R^4, all adapted except for a finite number of intermediate critical stages, at which one of the modifications described in Figure 10 takes place (between any two such critical stages, we have isotopy of diagrams in R^3, preserving ribbon intersections);

(b) *ribbon intersection sliding*, allowing a ribbon intersection to run along a 1-handle from one 0-handle to another one, as shown in Figure 11;

(c) *embedded 0/1-handles operations*, that is addition/delection of cancelling pairs of 0/1-handles and embedded 1-handle slidings (see Figure 12).

We observe that the second modification of Figure 10 is actually redundant in presence of the handle operations of Figure 12 (cf. proof of Proposition 1.3).
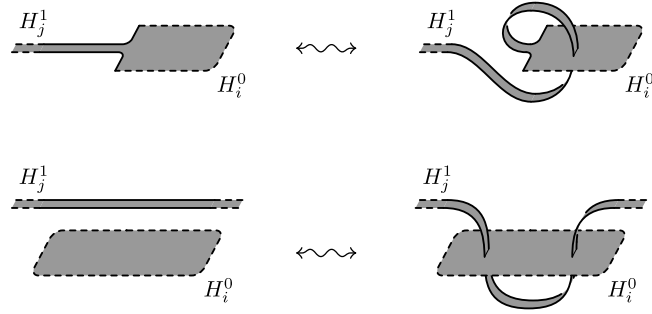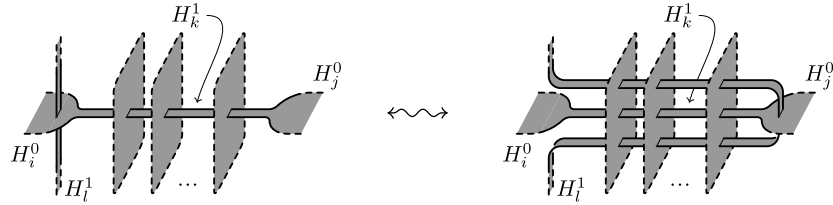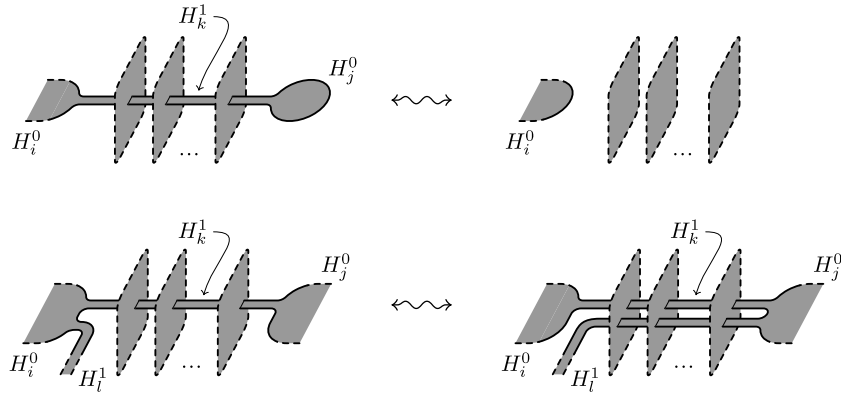
FIGURE 10.

FIGURE 11.

FIGURE 12.

It is also worth noticing that no twist appears in the 1-handle H_k^1 of Figures 11 and 12, since H_i^0 and H_j^0 can be assumed to be distinct in all the cases, up to addition/deletion of cancelling pairs of 0/1-handles, where they are always distinct.

PROPOSITION 1.2. *All the adapted 1-handlebody decompositions of a given ribbon surface are 1-equivalent as embedded 2-dimensional 1-handlebodies. More precisely, they are related to each other by the special cases without vertical disks of the moves of Figures 11 and 12, realized (up to isotopy of diagrams) in such a way that the surface is kept fixed.*

Proof. First of all, we observe that the moves specified in the statement allow us to realize the following two modifications: 1) split a 0-handle along any regular arc avoiding ribbon intersections in the diagram, into two 0-handles joined by a new 1-handle; 2) split a 1-handle at any transversal arc avoiding ribbon intersections in the diagram, into two 1-handles, by inserting a new 0-handle along it. We leave the straightforward verification of this to the reader.

Let $F = H_1^0 \cup \ldots \cup H_m^0 \cup H_1^1 \cup \ldots \cup H_n^1 = \bar{H}_1^0 \cup \ldots \cup \bar{H}_{\bar{m}}^0 \cup \bar{H}_1^1 \cup \ldots \cup \bar{H}_{\bar{n}}^1$ be any two 1-handlebody decompositions of a ribbon surface F, which we denote respectively by H and $\bar{H}$. After having suitably split the 1-handles, we can assume that any 1-handle contains at most one ribbon self-intersection of F and that this coincides with its cocore. Up to isotopy, we can also assume that the 1-handles of H and $\bar{H}$ whose cocore is the same self-intersection arc coincide. Let $H_1 = \bar{H}_1, \ldots,$ $H_k = \bar{H}_k$ be these 1-handles. Then, it suffices to see how to make the remaining 1-handles $H_{k+1}^1, \ldots, H_n^1$ into $\bar{H}_{k+1}^1, \ldots, \bar{H}_{\bar{n}}^1$, without changing $H_1^1, \ldots, H_k^1$.

Calling η_i (resp. $\bar{\eta}_j$) the cocore of H_i^1 (resp. $\bar{H}_j^1$), we have $\eta_1 = \bar{\eta}_1, \ldots, \eta_k = \bar{\eta}_k$, while the arcs $\eta_{k+1}, \ldots, \eta_n$ can be assumed to be transversal with respect to the arcs $\bar{\eta}_{k+1}, \ldots, \bar{\eta}_{\bar{n}}$. Up to isotopy, we can think of each 1-handle as a tiny regular neighborhood of its cocore, so that the intersection between $H_{k+1}^1 \cup \ldots \cup H_n^1$ and $\bar{H}_{k+1}^1 \cup \ldots \cup \bar{H}_{\bar{n}}^1$ consists only of a certain number h of small four-sided regions.

We eliminate all these intersection regions in turn, by pushing them outside F along the $\bar{H}_j^1$'s. This is done by performing on H moves of the types specified in the statement, as suggested by the following Figure 13, which concerns the l-th elimination. Namely, in (a) we assume that the intersection is the first one along $\bar{\eta}_j$ starting from $\mathrm{Bd}\, F$, then we generate the new 1-handle H_{n+l}^1 by 0-handle splitting to get (b), finally (c) is obtained by handle sliding.
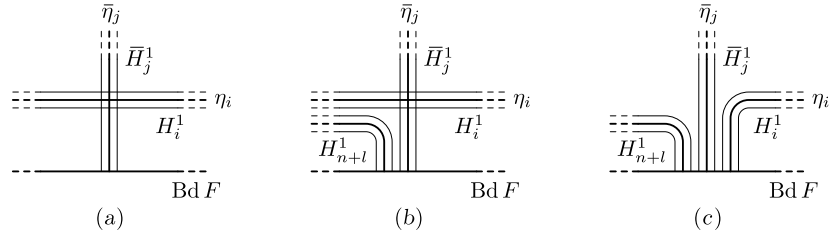

FIGURE 13.

After that, H has been changed into a new handlebody decomposition H' with 1-handles $H_1^1, \ldots, H_{n+h}^1$, such that H_i^1 is the same as above for $i \leq k$, while it is disjoint from the $\bar{H}_j^1$'s for $i > k$. Hence, $H_1^1, \ldots, H_k^1, H_{k+1}^1, \ldots, H_{n+h}^1, \bar{H}_{k+1}^1, \ldots, \bar{H}_{\bar{n}}^1$ can be considered as the 1-handles of a handlebody decomposition of F which can be obtained from both H' and $\bar{H}$ by 0-handle splitting. $\square$

Now, forgetting the 1-handlebody structure, 1-equivalence of embedded 2-dimensional 1-handlebodies induces an equivalence relation between ribbon surfaces, that we call *1-isotopy*. More precisely, two ribbon surfaces are 1-isotopic if and only if they admit 1-equivalent 1-handlebody decompositions. By the above proposition, this implies that actually all their 1-handlebody decompositions are 1-equivalent.

Of course 1-isotopy implies isotopy, but the converse is not known. In fact, the problem of finding a complete set of moves representing isotopy of ribbon surfaces is still open. We will come back to this delicate aspect later.

As we anticipated in the Introduction, the next proposition says that 1-isotopy is generated by the local isotopy moves of Figure 3, up to diagram isotopy in R^3, that means isotopy preserving ribbon intersections.

PROPOSITION 1.3. *Two ribbon surfaces are 1-isotopic if and only if they can be related by a finite sequence of diagram isotopies and moves $I_1, \ldots, I_4$ and their inverses.*

Proof. On one hand, we have to realize the modifications of Figures 10, 11 and 12, disregarding the handlebody structure, by moves $I_1, \ldots, I_4$ and their inverses. Of course, it is enough to do that in one direction, say from left to right. Proceeding in the order: one move I_1 suffices for the upper part of Figure 10, while the lower part can be obtained by combining one move I_2 with one move I_3; Figure 11 requires three moves for each vertical disk, one I_2, one I_3 and one I_4; the upper (resp. lower) part of Figure 12 can be achieved by one move I_2 (resp. I_3) for each vertical disk.

On the other hand, the surfaces of Figure 3 can be easily provided with adapted handlebody decompositions, in such a way that the relations just described between moves $I_1, \ldots, I_4$ and the above modifications can be reversed. In fact, only the special cases of those modifications with one vertical disk are needed. □

Branched coverings

A non-degenerate PL map $p : M \to N$ between compact PL manifolds of the same dimension m is called a *branched covering* if there exists an $(m-2)$-dimensional subcomplex $B_p \subset N$, the *branching set* of p, such that the restriction $p_| : M - p^{-1}(B_p) \to N - B_p$ is an ordinary covering of finite degree d. If B_p is minimal with respect to such property, then we have $B_p = p(S_p)$, where S_p is the *singular set* of p, that is the set of points at which p is not locally injective. In this case, both B_p and S_p, as well as the *pseudo-singular set* $S'_p = \mathrm{Cl}(p^{-1}(B_p) - S_p)$, are (possibly empty) homogeneously $(m-2)$-dimensional complexes.

Since p is completely determined, up to PL homeomorphisms, by the ordinary covering $p_|$ (cf. [11]), we can describe it in terms of its branching set B_p and its *monodromy* $\omega_p : \pi_1(N - B_p, *) \to \Sigma_d$, defined up to conjugation in Σ_d, depending on the choice of the base point $*$ and on the numbering of $p^{-1}(*)$. In particular, the monodromies of the meridians around the $(m-2)$-simplices of B_p determine the structure of the singularities of p. If all such monodromies are transpositions, then we say that p is *simple*. In this case, every point in the interior of a $(m-2)$-simplex of B_p is the image of one singular point, at which p is topologically equivalent to the complex map $z \mapsto z^2$, and $d-2$ pseudo-singular points.

Starting from $B_p \subset N$ and ω_p, we can explicitly reconstruct M and p by following steps: 1) choose a $(m-1)$-dimensional *splitting complex*, that means a subcomplex $C \subset N - \{*\}$ such that $B_p \subset C$ and the restriction $\omega_{p|} : \pi_1(N - C, *) \to \Sigma_d$ vanishes; 2) cut N along C in such a way that each $(m-1)$-simplex σ of C gives raise to 2 simplices σ^- and σ^+; 3) take d copies of the obtained complex (called the *sheets* of the covering) and denote by $\sigma_1^\pm, \ldots, \sigma_d^\pm$ the corresponding copies of $\sigma^\pm$; 4) identify in pairs the $\sigma_i^\pm$'s according to the monodromy $\rho = \omega_p(\alpha)$ of a loop α meeting C transversally at one point of σ, namely identify σ_i^- with $\sigma_{\rho(i)}^+$. Up to PL homeomorphisms, M is the result of such identification and p is the map induced by the natural projection of the sheets onto N.

A convenient representation of p can be given by labelling each $(m-2)$-simplex of B_p by the monodromy of a preferred meridian around it and each generator (in a finite generating set) of $\pi_1(N, *)$ by its monodromy, since those loops together

generate $\pi_1(N - B_p, *)$. Of course, only the labels on B_p are needed if N is simply connected. In any case, with a slight abuse of language if N is not simply connected, we refer to such a representation as a *labelled branching set*.

Two branched coverings $p : M \to N$ and $p' : M' \to N$ are called *equivalent* iff there exists PL homeomorphism $h : N \to N$ isotopic to the identity which lifts to a PL homeomorphism $k : M \to M'$. By the classical theory of ordinary coverings and [11], such a lifting k of h exists iff $h(B_p) = B_{p'}$ and $\omega_{p'} h_* = \omega_p$ up to conjugation in Σ_d, where $h_* : \pi_1(N - B_p, *) \to \pi_1(N - B_{p'}, h(*))$ is the homomophism induced by h. Therefore, in terms of labelled branching set, the equivalence of branched coverings can be represented by *labelled isotopy*.

By a *covering move*, we mean any non-isotopic modification making a labelled branching set representing a branched covering $p : M \to N$ into one representing a different branched covering $p' : M \to N$ between the same manifolds (up to PL homeomorphisms). We call such a move *local*, if the modification takes place inside a cell and can be performed whatever is the rest of labelled branching set outside. In the figures depicting local moves, we will draw only the portion of the labelled branching set inside the relevant cell, assuming everything else to be fixed.

As a primary source of covering moves, we consider the following two very general equivalence principles (cf. [39]). Several special cases of these principles have already appeared in the literature and we can think of them as belonging to the "folklore" of branched coverings.

DISJOINT MONODROMIES CROSSING. *Subcomplexes of the branching set of a covering that are labelled with disjoint permutations can be isotoped independently from each other without changing the covering manifold.*

The reason why this principle holds is quite simple. Namely, being the labelling of the subcomplexes disjoint, the sheets non-trivially involved by them do not interact, at least over the region where the isotopy takes place. Hence, the relative position of such subcomplexes is not relevant in determining the covering manifold. Typical applications of this principle are the local moves M_2 and R_2 (cf. Figures 1 and 2).

It is worth observing that, abandoning transversality, the disjoint monodromies crossing principle also gives the special case of the next principle when the σ_i's are disjoint and L is empty.

COHERENT MONODROMIES MERGING. *Let $p : M \to N$ be any branched covering with branching set B_p and let $\pi : E \to K$ be a connected disk bundle imbedded in N, in such a way that: 1) there exists a (possibly empty) subcomplex $L \subset K$ for which $B_p \cap \pi^{-1}(L) = L$ and the restriction of π to $B_p \cap \pi^{-1}(K - L)$ is an unbranched covering of $K - L$; 2) the monodromies $\sigma_1, \ldots, \sigma_n$ relative to a fundamental system $\omega_1, \ldots, \omega_n$ for the restriction of p over a given disk $D = \pi^{-1}(x)$, with $x \in K - L$, are coherent in the sense that $p^{-1}(D)$ is a disjoint union of disks. Then, by contracting the bundle E fiberwise to K, we get a new branched covering $p' : M \to N$, whose branching set $B_{p'}$ is equivalent to B_p, except for the replacement of $B_p \cap \pi^{-1}(K - L)$ by $K - L$, with the labelling uniquely defined by letting the monodromy of the meridian $\omega = \omega_1 \ldots \omega_n$ be $\sigma = \sigma_1 \ldots \sigma_n$.*

We remark that, by connectedness and property 1, the coherence condition required in 2 actually holds for any $x \in K$. Then, we can prove that p and p' have the

same covering manifold, by a straightforward fiberwise application of the Alexander's trick to the components of the bundle $\pi \circ p : p^{-1}(E) \to K$. A coherence criterion can be immediately derived from Section 1 of [35].

The coherent monodromy merging principle originated from a classical perturbation argument in algebraic geometry and appeared in the literature as a way to deform non-simple coverings between surfaces into simple ones, by going in the opposite direction from p' to p (cf. [5]). In the same way, it can be used in dimension 3, both for achieving simplicity (cf. [16]) and removing singularities from the branching set. We will do that in the proof of Theorem 4 by means of the moves S_1 and S_2 of Figure 7, which are straightforward applications of this principle. Actually, analogous resuts could be proved in dimension 4, but we will not do it here.

The coherent monodromy merging principle, also provides an easy way to verify that M_1 and R_1 are local covering moves, as shown in Figures 14 and 15. In both these figures, we apply the principle for going from (a) to (b) and from (c) to (d), while (b) and (c) are equivalent up to labelled isotopy.
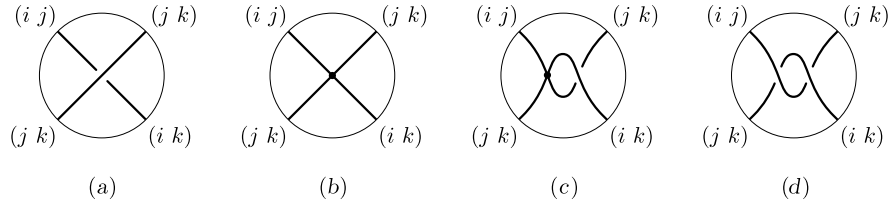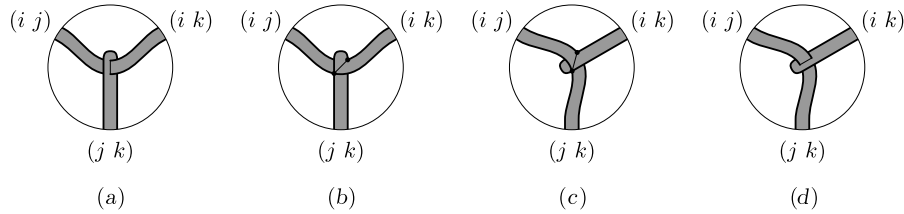

FIGURE 14.

FIGURE 15.

So far we have seen that all the moves presented in the Introduction, except for moves T and $P_{\pm}$, are local covering moves. However, we will give a different proof of that for moves R_1 and R_2 in Section 2, by relating them to 2-deformations of 4-handlebodies.

Now, we consider the notion of stabilization that appears in all the equivalence theorems stated in the Introduction. This is a particular local covering move, which makes sense only for branched coverings of S^m or B^m and, differently from all the previous moves, changes the degree of the covering, increasing it by one.

STABILIZATION. *A branched covering $p : M \to S^m$ (resp. $p : M \to B^m$) of degree d, can be stabilized to degree $d + 1$ by adding to the labelled branching set a trivial separate $(m - 2)$-sphere (resp. regularly embedded $(m - 2)$-disk) labelled with the transposition $(i\ d{+}1)$, for some $i = 1, \ldots, d$.*

The covering manifold of such a stabilization is still M, up to PL homeomorphisms. In fact, it turns out to be the connected sum (resp. boundary connected sum) of M itself, consisting of the sheets $1, \ldots, d$, with the copy of S^m (resp. B^m) given by the extra trivial sheet $d + 1$.

By *stabilization to degree n* (or *n-stabilization*) of a branched covering $p : M \rightarrow S^m$ (resp. $p : M \rightarrow B^m$) of degree $d \leq n$ we mean the branched covering of degree n obtained from it by performing $n - d$ stabilizations as above. In particular, this leaves p unchanged if $d = n$.

We conclude this paragraph by focusing on the branched coverings we will deal with in the following sections, that is coverings of S^3 branched over links or embedded graphs and coverings of B^4 branched over ribbon surfaces. We recall that in this context PL and smooth are interchangeable.

We represent a d-fold covering of $p : M \rightarrow S^3$ branched over a link $L \subset S^3$, by a Σ_d-labelled oriented diagram D of L describing the monodromy of p in terms of the Wirtinger presentation of $\pi_1(S^3 - L)$ associated to D. Namely, we label each arc of D by the monodromy of the standard positive meridian around it. Of course, the Wirtinger relations impose constraints on the labelling at crossings, and each Σ_d-labelling of D satisfying such constraints do actually represent a d-fold covering of S^3 branched over L. Then, labelled isotopy can be realized by means of labelled Reidemeister moves.

For simple coverings, the orientation of D is clearly unnecessary and there are three possible ways of labelling the arcs at each crossing: either all with the same transposition $(i\ j)$ or like at the two crossings in the left side of Figure 1.

The Montesinos-Hilden-Hirsch representation theorem of closed connected oriented 3-manifolds as branched coverings of S^3 (see Introduction), can be formulated in terms of labelled link diagrams, with labels taken from the three transpositions of Σ_3, according to the above labelling rules at crossings.

The extension from branching links to branching embedded graphs is straightforward. In fact, we only need to take into account extra labelling constraints and labelled moves at the vertices of the graph.

Finally, let us consider a d-fold covering $p : M \rightarrow B^4$ branched over a ribbon surface $F \subset B^4$. Again, we represent the monodromy in terms of the Wirtinger presentation of $\pi_1(B^4 - F)$ associated to a locally oriented diagram of F. Actually, since we will only consider simple coverings, we will never need local orientations.

The same labelling rules as above apply to ribbon intersections (cf. Figure 2) as well as to ribbon crossings. However, contrary to what happens for ribbon intersections, when a ribbon crosses under another one, its label changes only locally (at the undercrossing region). We notice that, if $F \subset B^4$ is a labelled ribbon surface representing a d-fold (simple) covering of $p : M \rightarrow B^4$, then $L = F \cap S^3$ is a labelled link representing the restriction $p_{|\mathrm{Bd}} : \mathrm{Bd}\, M \rightarrow S^3$. This is still a d-fold (simple) covering, having the diagram of F as a splitting complex.

As mentioned in the Introduction, labelled ribbon surfaces in B^4 (that is coverings of B^4 branched over ribbon surfaces) represent all the 4-dimensional 2-handlebodies. By Montesinos [31] (cf. next Section 2), for the connected case it suffices to take labels from the three transpositions of Σ_3 (that is to consider 3-fold simple coverings).

Though labelled isotopy of branching ribbon surfaces preserves the covering manifold M up to PL homeomorphisms, we are interested in the (perhaps more restrictive) notion of *labelled 1-isotopy*, which preserves M up to 2-deformations (cf. Lemma 2.3). This can be realized by means of labelled diagram isotopy and labelled 1-isotopy moves, that is diagram isotopy and 1-isotopy moves of Figure 3, suitably labelled according to the above rules.

Kirby diagrams

A *Kirby diagram* describes an orientable 4-dimensional 2-handlebody $H^0 \cup H_1^1 \cup \ldots \cup H_m^1 \cup H_1^2 \cup \ldots \cup H_n^2$ with only one 0-handle, by encoding 1- and 2-handles in a suitable link $K \subset S^3 \cong \mathrm{Bd}\, H^0$. Namely, K has m dotted components spanning disjoint flat disks which represent the 1-handles and n framed components which determine the attaching maps of the 2-handles. We refer to [23] or [14] for details and basic facts about Kirby diagrams, limiting ourselves to recall here only the relevant ones for our purposes.

The assumption of having only one 0-handle is not so restrictive. In fact, given any connected handlebody, the union of 0- and 1-handles contracts in a natural way to a connected graph G. Then, by choosing a maximal tree $T \subset G$ and fusing all the 0-handles together with the 1-handles corresponding to the edges of T, we get a new handlebody with only one 0-handle. This fusion process can be performed by 1-handle slidings and 0/1-handle cancellation, so the new handlebody is equivalent to the original one. As a consequence, different choises of the tree T give raise to handlebodies which are equivalent up 1-handle sliding. This fact immediately implies that k-equivalence between handlebodies having only one 0-handle can be realized without adding any extra 0-handle.

On the other hand, the same assumption of having only one 0-handle, is crucial in order to make a natural convention on the framings, that allows to express them by integers fixing as zero the homologically trivial ones.

However, at least in the present context, it seems preferable to renounce this advantage on the notation for framings in favour of more flexibility in the representation of multiple 0-handles. The reason is that a d-fold covering of B^4 branched over a ribbon surface (actually an embedded 2-dimensional 1-handlebody) turns out to have a natural handlebody structure with d 0-handles.

Of course, the reduction to only one 0-handle is still possible but it must be performed explicitly. This makes the connection between branched coverings and ordinary Kirby diagrams more clear and transparent than before.

We call a generalized Kirby diagram our representation of an orientable 4-dimensional 2-handlebody with multiple 0-handles. It is essentially defined by overlapping the boundaries of all the 0-handles to let the diagram take place in S^3 and by putting labels in the diagram in order to keep trace of the original 0-handle where each part of it is from. If there is only one 0-handle, the labels can be omitted and we have an ordinary Kirby diagram.

More precisely, a *generalized Kirby diagram* representing an orientable 4-dimensional handledoby $H_1^0 \cup \ldots \cup H_d^0 \cup H_1^1 \cup \ldots \cup H_m^1 \cup H_1^2 \cup \ldots \cup H_n^2$ consists of the following data: a boxed label indicating the number d of 0-handles; m dotted unknots spanning disjoint flat disks, each side of which has a label from $\{1, \ldots, d\}$;

n framed disjoint knots transversal with respect to those disks, with a label from $\{1, \ldots, d\}$ for each component of the complement of the intersections with the disks. The labelling must be admissible in the sense that all the framed arcs coming out from one side of a disks have the same label of that side (cf. Figure 16). This rule makes the labelling redundant and some times we will omit the superfluous labels. Moreover, being uniquely related to the indexing of the 0-handles, the labelling must be considered defined up to permutation of $\{1, \ldots, d\}$. Finally, the framings are always drawn as parallel curves, hence no confusion arises with labels.
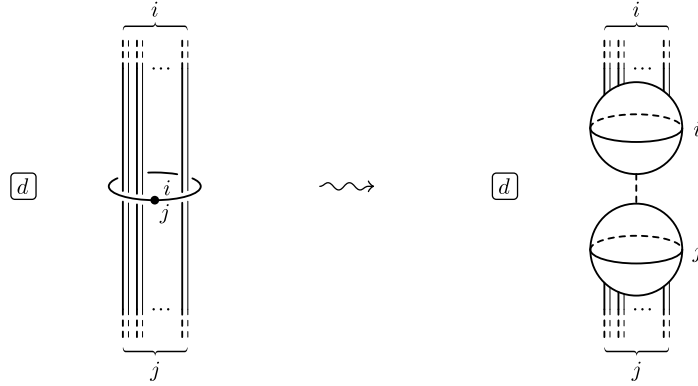
FIGURE 16.

To establish the relation between a generalized Kirby diagram and the handle-body it represents, we first convert dot notation for 1-handles into ball notation, as shown in Figure 16. Here, the two balls, together with the relative framed arcs, are symmetric with respect to the horizontal plane containing the disk and squeezing them vertically on the disk we get back the original diagram. After that, we consider the disjoint union of 0-handles $H_1^0 \cup \ldots \cup H_d^0$ and draw on the boundary of each H_i^0 the portion of the diagram labelled with i, no matter how we identify such boundary with S^3. Then, we attach to $H_1^0 \cup \ldots \cup H_d^0$ a 1-handle between each two paired balls (possibly lying in different 0-handles), according to the diffeomorphism induced by the above symmetry, so that we can join longitudinally along the handle the corresponding framed arcs. Of course, the result turns out to be defined only up to 1-handle full twists. At this point, we have a 1-handlebody $H_1^0 \cup \ldots \cup H_d^0 \cup H_1^1 \cup \ldots \cup H_m^1$ with n framed loops in its boundary and we use such framed loops as attaching instructions for the 2-handles $H_1^2, \ldots, H_n^2$.

We observe that any orientable 4-dimensional 2-handlebody can be represented, up to isotopy, by a generalized Kirby diagram. In fact, in order to reverse our construction, we only need that the identification of the boundaries of the 0-handles with S^3 is injective on the attaching regions of 1- and 2-handles and that the attaching maps of the 2-handles run longitudinally along the 1-handles. These properties can be easily achieved by isotopy.

Sometimes, it will be convenient to derogate from the prescribed labelling rule for generalized Kirby diagrams, by allowing a framed component with label k to cross a disk spanned by a dotted component with labels i and j, provided that $k \notin \{i, j\}$. Clearly, such a crossing does not mean that the framed loop goes over the 1-handle
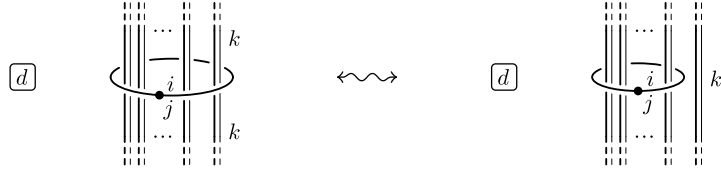
FIGURE 17.

corresponding to the dotted one, since it originates from the identification of different 0-handles. Figure 17 depictes the way to eliminate it.

The above construction gives isotopic handlebody structures if and only if the starting generalized Kirby diagrams are equivalent up to *labelled isotopy*, generated by labelled diagram isotopy, preserving all the intersections between loops and disks (as well as labels), and by the three moves described in Figure 18. Here, we assume $k \neq l$, so that the crossing change at the bottom of the figure preserves the isotopy class of the framed link in $H_1^0 \cup \ldots \cup H_d^0 \cup H_1^1 \cup \ldots \cup H_m^1$. It is worth remarking that, due to this crossing change, the framing convention usually adopted for ordinary Kirby diagrams cannot be extended to generalized Kirby diagrams.
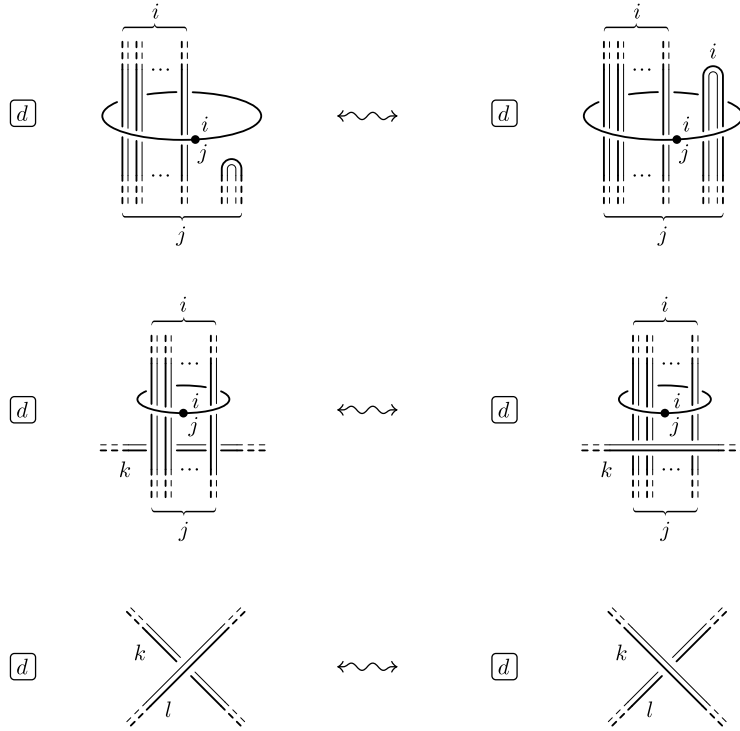
FIGURE 18.

On the contrary, the other two moves make sense whatever are i, j and k. In particular, if $i = j = k$ they reduce to the ordinary ones. Actually, this is the only relevant case for the second move, usually referred to as "sliding a 2-handle over a 1-handle", being the other cases obtainable by crossing changes. Moreover, even this ordinary case becomes superfluous in the context of 2-deformations, since it can be realized by addition/deletion of cancelling 1/2-handles and 2-handle sliding (cf. [14]).

The following Figures 19 and 20 show how to represent 2-deformations of 4-dimensional 2-handlebodies in terms of generalized Kirby diagrams. Namely, the moves of Figure 19 correspond to addition/deletion of cancelling 0/1-handles (on the right side we assume $i \leq d$) and 1/2-handles, while the moves of Figure 20 correspond to 1- and 2-handle sliding. Except for the addition/deletion of cancelling 0/1-handles, which does not make sense for ordinary Kirby diagrams, also the rest of the moves reduce to the ordinary ones if $i = j = k$.
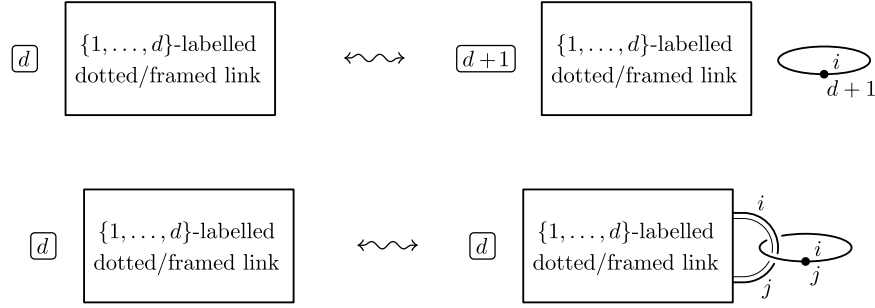
FIGURE 19.
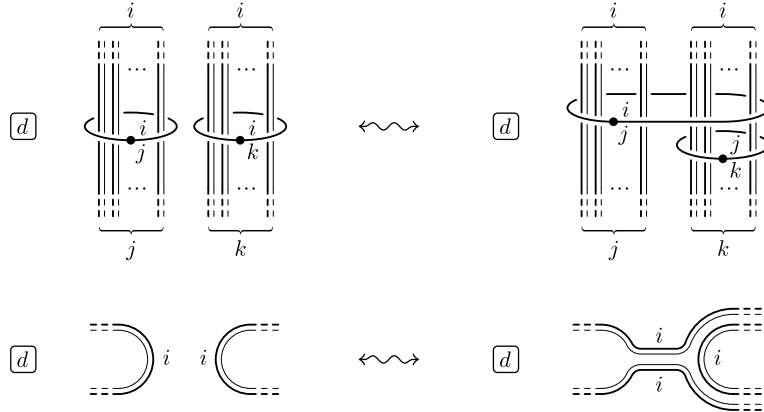
FIGURE 20.

The 1-handle sliding is included for the sake of completeness, but it can be generated by addition/deletion of cancelling 1/2-handles and 2-handle sliding, just like in the ordinary case (cf. [14]).

Summing up, two generalized Kirby diagrams represent 2-equivalent 4-dimensional 2-handlebodies if and only if they can be related by the first and third moves of Figure 18 (labelled isotopy), the two moves of Figure 19 (addition/deletion of cancelling handles) and the second move of Figure 20 (2-handle sliding).

Of these, only the first move of Figure 18 and the second ones of Figures 19 and 20 (for $i = j = d = 1$) make sense in the case of ordinary Kirby diagrams. Actually, such three moves suffice to realize 2-equivalence of 4-dimensional 2-handlebodies with only one 0-handle, since any extra 0-handle occurring during a 2-deformation can be eliminated by a suitable fusion of 0-handles.

The main theorem of Kirby calculus [22] asserts that two orientable 4-dimensional 2-handlebodies have diffeomorphic boundaries if and only if they are related by 2-deformations, blowing up/down and 1/2-handle trading.

In terms of generalized Kirby diagrams these last two modifications can be realized by the moves of Figure 21. These moves essentially coincide with the corresponding ones for ordinary Kirby diagrams (with $i = d = 1$), being the involved labels all the same.
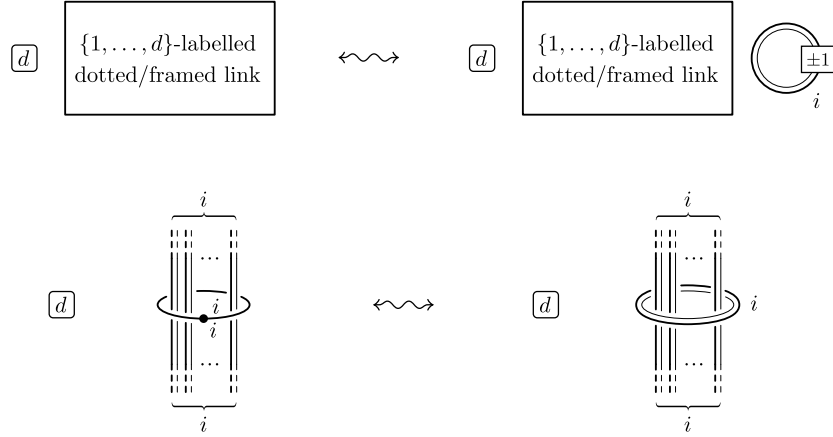

FIGURE 21.

We conclude this paragraph, by coming back to ordinary Kirby diagrams and in particular by introducing the standard form that will be used in Section 3.

FIGURE 22.

– 21 –
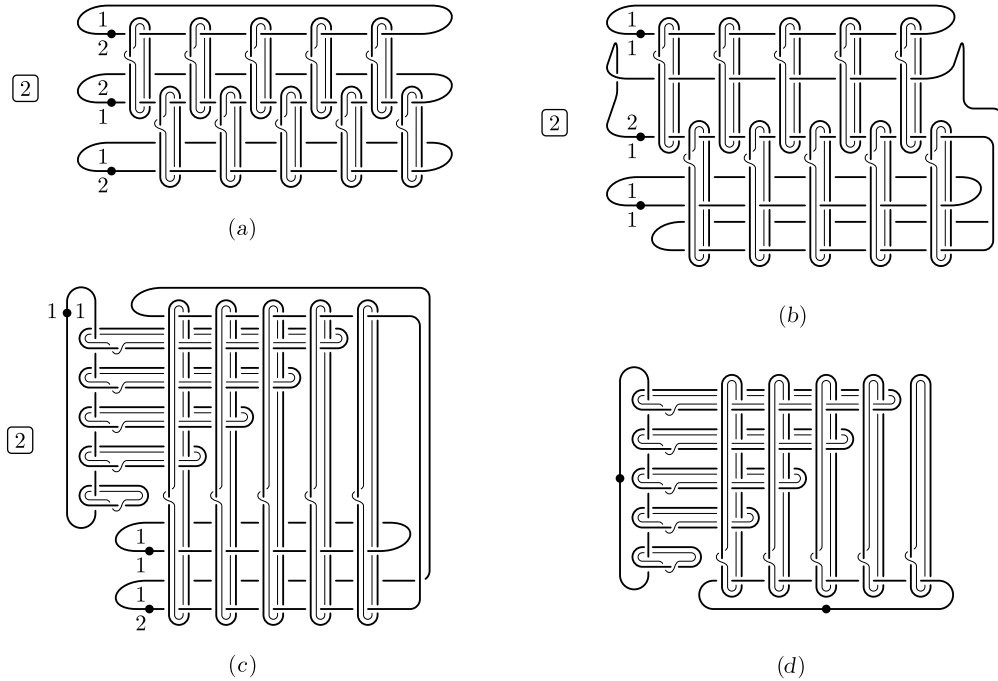
First let us observe that, given any generalized Kirby diagram representing a connected handlebody, we can use 2-deformation moves to transform it into an ordinary one, by reducing the number of 0-handles to 1. In fact, assuming $d > 1$, we can eliminate the d-th handle as follows (see Figure 22 for an example with $d = 2$): perform 1-handle sliding in order to leave only one label of one dotted unknot equal to d; untangle such unknot from the rest of the diagram by labelled isotopy; eliminate the d-th 0-handle by 0/1-handle cancellation.
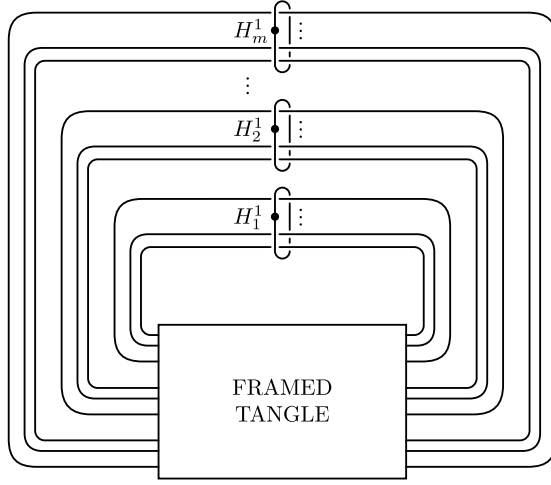
FIGURE 23.
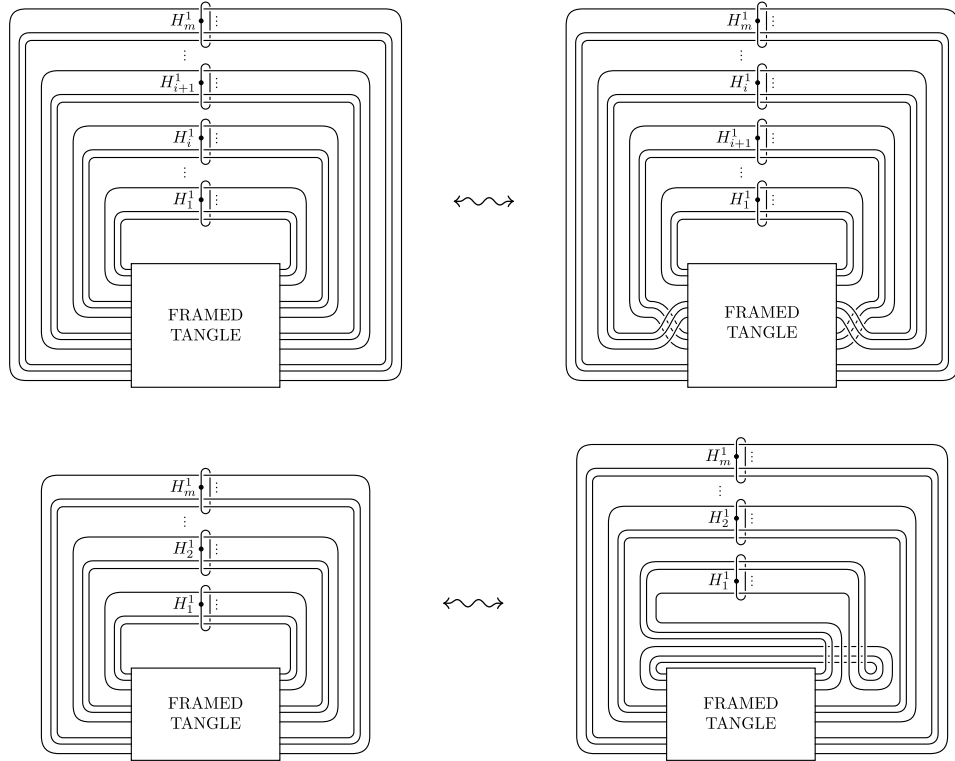
FIGURE 24.

An ordinary Kirby diagram is said to be in *standard form* if it looks like in Figure 23, where all the framings are understood to coincide with the blackboard one outside the box. Apparently, any ordinary Kirby diagram can be isotoped into such a standard form. Moreover, isotopy between Kirby diagrams in standard form consists of framed isotopy fixing the dotted disks together with the two moves of Figure 24, which relate different choises for the vertical order and the orientation of the 1-handles.

It is straightforward to see that, apart from the move of Figure 25 where an arc of the tangle is isotoped outside the box to pass between two the dotted disks H_i^1 and H_{i+1}^1, any isotopy between Kirby diagrams in standard form which fix the dotted disks can be assumed to move only the framed tangle inside the box.
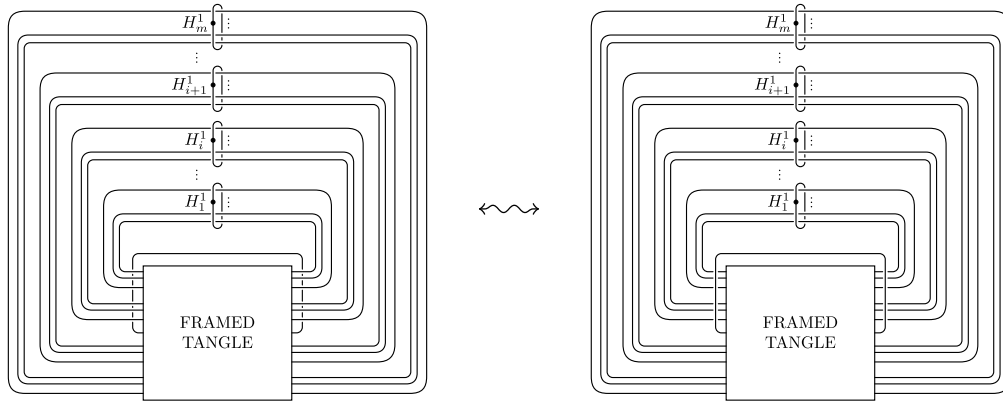
FIGURE 25.

Finally, in the following proposition, we recapitulate the moves relating ordinary Kirby diagrams which are 2-equivalent or have diffeomorphic boundaries.

PROPOSITION 1.4. *Given two ordinary Kirby diagrams K and K' in standard form, denote by H and H' the corresponding 4-dimensional 2-handlebodies. Then:*

(a) *H and H' are 2-equivalent if and only if K and K' are equivalent up to the first move of Figure 18, the second ones of Figures 19 and 20, the two moves of Figure 24 and framed tangle isotopy inside the box;*

(b) *H and H' have diffeomorphic boundaries if and only if K and K' are equivalent up to the moves listed in (a), the two moves of Figure 21 and framed tangle isotopy inside the box.*

All the moves are understood to preserve the standard form (up to diagram isotopy) in the obvious way and the ones of Figures 18, 19, 20 and 21 are considered only in the ordinary case, that is for $i = j = d = 1$.

Proof. It is enough to prove (a), being (b) an immediate consequence of it and of the main theorem of Kirby calculus. To do that, we only need to show that the moves listed in (a) allow us to represent any 2-deformation between H and H'.

As discussed above, 2-deformations of ordinary Kirby diagrams can be represented by the ordinary cases of the first move of Figure 18 and of the second ones of Figures 19 and 20, together with diagram isotopy. Moreover, up to diagram isotopy,

any of these moves can be performed on a Kirby diagram in standard form, in such a way that the standard form is preserved.

On the other hand, we already observed that the moves of Figures 24 and 25, together with framed tangle isotopy inside the box, allows us to realize diagram isotopy between Kirby diagrams in standard form.

Then, to conclude the proof, it suffices to notice that the move of Figure 25 can be obtained by sliding of a 2-handle over some 1-handles (cf. second move of Figure 18). Hence, in the context of 2-deformations diagram isotopy between Kirby diagrams in standard form can be realized without it. $\square$

2. From labelled ribbon surfaces to Kirby diagrams

The aim of this section is to show how any adapted 1-handlebody structure on a labelled ribbon surface F representing a d-fold simple branched covering $p : M \to B^4$ naturally induces a 2-handlebody structure on M defined up to 2-deformations.

In this context, naturally means that labelled embedded 1-deformations on F induce 2-deformations on M. Then, by Propositions 1.2 and 1.3, M turns out to be endowed with a 2-handlebody structure, whose 2-equivalence class is uniquely determined by the labelled 1-isotopy class of F. We denote by K_F the generalized Kirby diagram corresponding to such 4-dimensional 2-handlebody structure (defined up to 2-deformations).

Moreover, we will see that the 2-equivalence class of K_F is also preserved by the covering moves R_1 and R_2 of Figure 2 and we will discuss some consequences of this fact. In particular, we will introduce some auxiliary moves generated by R_1 and R_2, that will be needed in the next sections.

Let us start with the construction of K_F. Given a labelled ribbon surface F as above with an adapted 1-handlebody decomposition, we can write $F = D_1 \cup \ldots \cup D_m \cup B_1 \cup \ldots \cup B_n$, where the D_h's are disjoint flat disks (the 0-handles of F) while the B_h's are disjoint bands attached to $F_0 = D_1 \cup \ldots \cup D_m$ (the 1-handles of F). Looking at the diagram of F in R^3 and using for it the same notations as for F itself, we see that the D_h's, as well as the B_h's, are still disjoint from each other, while any band B_h may form ribbon intersections with the disks $D_1, \ldots, D_m$.

We denote by $p_0 : M_0 \to B^4$ the simple covering determined by the labelled surface $F_0 \subset B^4$. The covering manifold M_0 turns out to be a 4-dimensional handlebody with d 0-handles and a 1-handle H_h^1 for each disk D_h (cf. [31]). A generalized Kirby diagram of M_0 can be immediately obtained by replacing any disk D_h by a dotted unknot coinciding with its boundary, as shown in Figure 26. Here, there are two possible way to assign the labels i and j to the two faces of D_h. We call such an assignment a *polarization* of the disk D_h.

FIGURE 26.

Now, following [31] (cf. also [20]), we have that any band B_h attached to F_0 gives rise to a 2-handle H_h^2 attached to M_0 along the framed loop given by the unique annular component of $p_0^{-1}(B_h)$.

In order to describe a labelled framed loop representing H_h^2 in the generalized Kirby diagram, let us call D_{h_1} and D_{h_2} the (possibly coinciding) disks of F_0 at which B_h is attached. Disregarding for the moment the ribbon intersections of B_h with F_0, such framed loop is given by two parallel copies of B_h lying on opposite sides, joined together to form ribbon intersections with D_{h_1} and D_{h_2} and labelled consistently with the polarizations of those disks, as suggested by Figure 27.
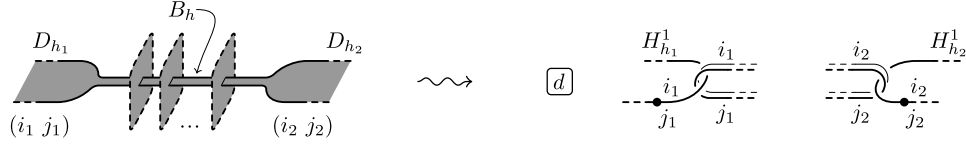
FIGURE 27.
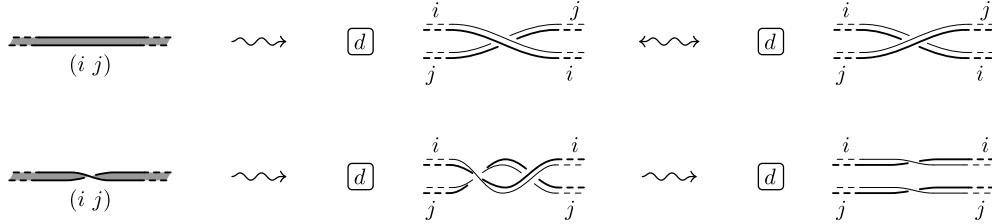
FIGURE 28.

Actually, to have simultanous labelling consistency at both ends of B_h, we may be forced to interchange the two copies of B_h by a crossover, as in the upper part of Figure 28. The two ways to realize the crossover are equivalent up to labelled isotopy, since $i \neq j$. Notice that we can perform crossovers wherever we want along B_h, provided their number has the right parity to respect labelling consistency. In the lower part of Figure 28 we see that, up to crossovers, twists along B_h contribute only to the framing and not to the isotopy type of the corresponding loop. Namely, each positive (resp. negative) half twist along B_h gives rise to a positive (resp. negative) full twist in the framing.

Figure 29 explains how to interpret a single ribbon intersection between B_h and F_0 into the generalized Kirby diagram, in the four possible cases depending on the monodromies associated to B_h and F_0 at that intersection. Here, we assume that i, j, k and l are all distinct and use the notation of Figure 17 for the intermediate steps. In all the cases, the construction is carried out in a regular neighborhood of an arc α contained in F_0 and joining the ribbon intersection with $\mathrm{Bd}\, F_0$. The labels of the two copies of B_h in the generalized Kirby diagram are determined by monodromies associated to B_h before and after the intersection and by the side from which α approaches B_h. In the first and third cases, we introduce a kink to allow labelling consistency of both the copies of B_h with respect to the disk (there are two different ways to realize such a kink, but they are equivalent up to labelled isotopy). To make all the local labellings at the ribbon intersections fit together with
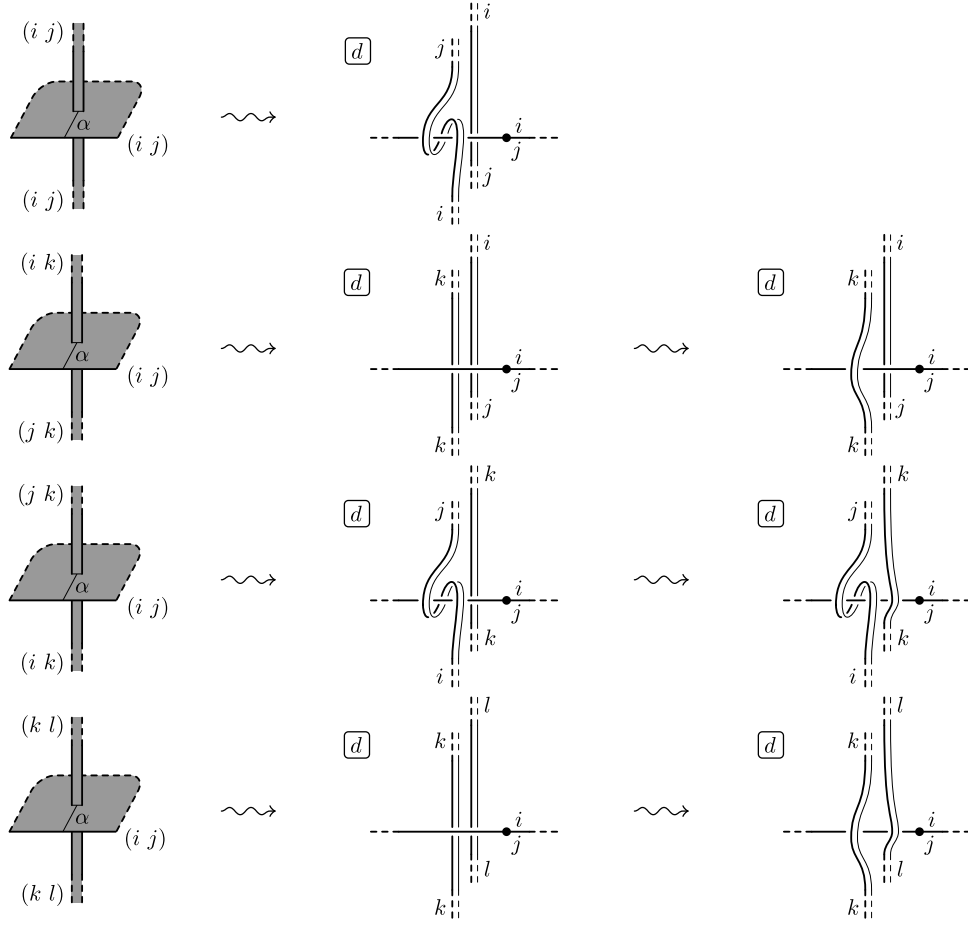
FIGURE 29.

each other along the two copies of B_h and with the ones already fixed at ends of B_h, we use again crossovers.

We conclude the definition of K_F, by specifying that the arcs α related to different ribbon intersections are assumed to be disjoint, in such a way that the corresponding constructions do not interact.

Our next aim is to show that K_F is well defined up to 2-deformation moves, in the sense that the 2-equivalence class of the corresponding 4-dimensional 2-handlebody depends only on the labelled ribbon surface F. As a preliminary step, we prove the following Lemma concerning the choices involved in the construction of K_F from a 1-handlebody structure of F.

LEMMA 2.1. *Let $F \subset B^4$ be a labelled ribbon surface representing a d-fold simple branched covering $p : M \to B^4$. Then, the generalized Kirby diagram K_F, constructed starting from a given adapted 1-handlebody structure on F, describes a 4-dimensional 2-handlebody structure on M, whatever choices we make for the polarizations, the crossovers and the arcs α. Moreover, such 4-dimensional 2-handlebody structure is uniquely determined up to handle isotopy.*

Proof. Let $F = D_1 \cup \ldots \cup D_m \cup B_1 \cup \ldots \cup B_n$ be an adapted 1-handlebody decomposition of F as in the definiton of K_F and let us adopt here all the notations related to it we introduced there.

Then, the lemma immediately follows from [31], once one has checked that the framed link of K_F does really represent, up to handle isotopy, the framed link in M_0 consisting of the unique annular component of $p_0^{-1}(B_h)$ for each band B_h of F. Taking into account what we have said above, this is a straightforward consequence of the very definition of generalized Kirby diagram.

Nevertheless, for the convenience of the reader, we skecth a direct proof of the independence of F_K, up to handle isotopy, on the choices involved in its construction.

We have already observed that crossovers are not relevant up to labelled isotopy. Concerning the arcs α, it suffices to prove that the elementary moves of Figure 30, where we replace a single arc α by α', preserve K_F up to 2-deformation moves. Simple inspection of all the cases confirms that once again only labelled isotopy moves are needed.
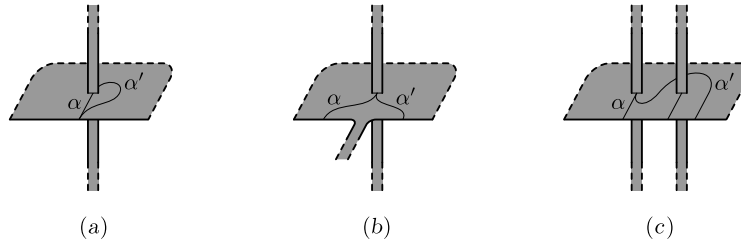

(a)   (b)   (c)

FIGURE 30.

Thus, it remains to see what happens when we invert the polarization of a disk D_h. The relative dotted unknot with the different labellings giving the two possible polarizations of D_h is drawn in Figure 31 (a) and (d). Here, we assume that
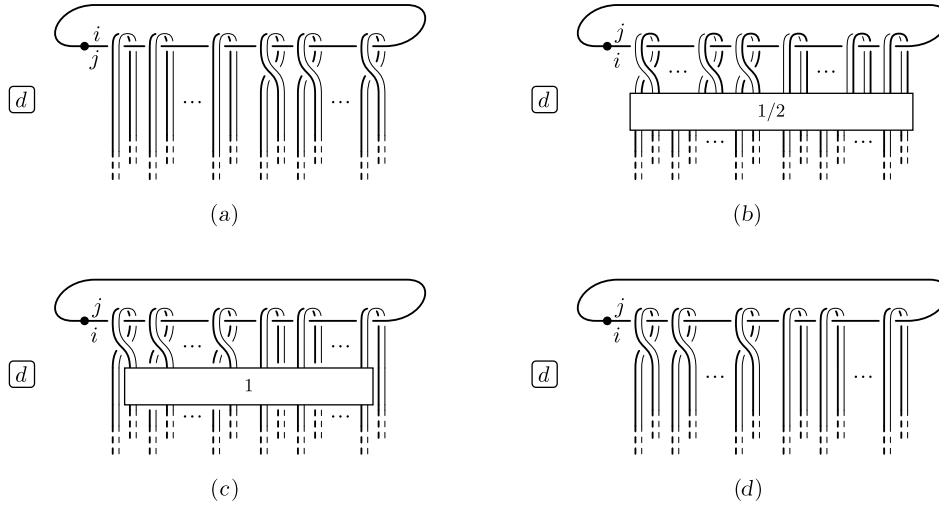

(a)   (b)

(c)   (d)

FIGURE 31.

the framed arcs passing thought D_h, coming either from bands attached to D_h or from ribbon intersection of bands with D_h, have been isotoped all together into a canonical position. For the sake of clarity, we sorted such labelled arcs to separate the ones which have been kinked for respecting labelling consistency. To see that the diagrams (a) and (d) of Figure 31 are equivalent up to 2-deformation moves, we

consider the other ones as intermediate steps. We start by isotoping upside down the dotted unknot of (a) to obtain (b). Then, we use labelled isotopy once again to make the arcs labelled by i and the ones labelled by j form separate positive half twists. These two half twists add up to give a unique positive full twist in (c). Finally, we get (d) by performing a negative twist on the 1-handle represented by the dotted unknot. Such a 1-handle twist can be easily realized by the second labelled isotopy move of Figure 18. $\square$

PROPOSITION 2.2. *Let $F \subset B^4$ be a labelled ribbon surface representing a d-fold simple branched covering $p : M \to B^4$. Then, the generalized Kirby diagrams K_F constructed starting from different adapted 1-handlebody structures on F, describe 2-equivalent 4-dimensional 2-handlebody structures on M. That is, K_F is uniquely determined by F up to 2-deformation moves.*

Proof. We observe that any labelled diagram isotopy (preserving ribbon intersections) on F induces a labelled isotopy on K_F as a generalized Kirby diagram. Hence, the statement follows from Proposition 1.2, once we prove that performing on F labelled versions of the moves of Figures 11 and 12 without vertical disks corresponds to modifying K_F by certain 2-deformation moves.

In all the cases we can choose the same polarization for H_i^0 and H_j^0, since these can be assumed to be distinct 0-handles (cf. notice after Figure 12). Then, apparently the two moves of Figure 12 correspond respectively to addition/deletion of a cancelling pair of 1/2-handles and to sliding the 2-handle deriving from H_l^1 over the one deriving from H_k^1. Similarly, in the case of move of Figure 11 we have two slidings involving the same 2-handles, one sliding for each of the two parallel copies of H_l^1 forming the framed loop originated from it. We leave to the reader the straightforward verification of this fact for all the four cases of Figure 29. $\square$

A very simple example of the above construction, without ribbon intersections, is depicted in Figure 32. Here, the adapted 1-handlebody structure of the labelled ribbon surface on the left is the obvious one with 3 horizontal 0-handles and 10 vertical 1-handles, while the resulting generalized Kirby diagram on the right is the same of Figure 22. We notice that, for a double covering of B^4 branched over ribbon surface without ribbon intersections, such as the one of Figure 32, the handlebody presentation we obtain by our construction coincides, after suitable reduction to ordinary Kirby diagram, with the one given in [2].
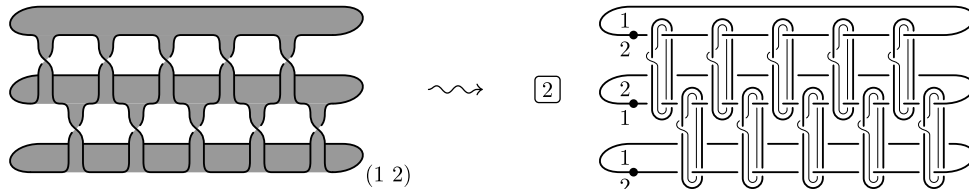

FIGURE 32.

The following Proposition 2.4 tells us that the 2-equivalence class of K_F actually depends only on the labelled 1-isotopy class of F and it is also preserved by stabilization and covering moves R_1 and R_2. This is essentially the "only if" part of Theorem 1.

LEMMA 2.3. *If the labelled ribbon surfaces $F, F' \subset B^4$, representing d-fold simple branched coverings of B^4, are related by labelled 1-isotopy, then the generalized Kirby diagrams K_F and $K_{F'}$ are equivalent up to 2-deformation moves.*

Proof. By Proposition 1.3, labelled 1-isotopy is generated by labelled diagram isotopy and the labelled versions of moves $I_1, \ldots, I_4$ (cf. Figure 3). Since labelled diagram isotopy on F induces labelled isotopy on K_F as a generalized Kirby diagram, we have only to deal with the moves.

Move I_1 admits a unique labelling up to conjugation in Σ_d. Generalized Kirby diagrams arising from the labelled ribbon surfaces involved in the resulting labelled move are depicted in Figure 33 (we assume the surfaces endowed with the handlebody structures of the corresponding move of Figure 10). As the reader can easily check, such diagrams are related by labelled isotopy.
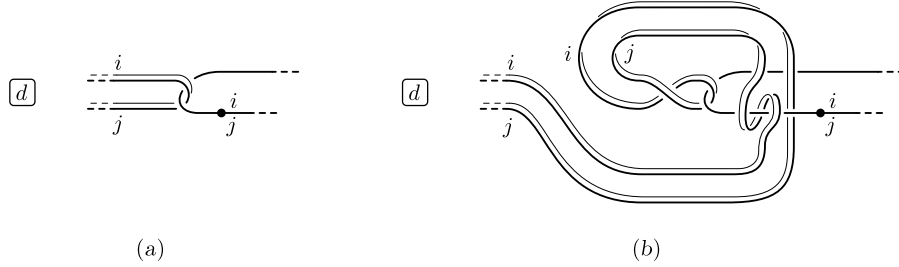

(a)   (b)

FIGURE 33.

Moves I_2 and I_3 admit three distinct labellings up to conjugation in Σ_d. Namely, if $(i\ j)$ is the label of the horizontal component, then the top end of the vertical one can be labelled by $(i\ j)$, $(j\ k)$ or $(k\ l)$.

The first case is considered in Figure 34 for I_2 and Figure 35 for I_3. Looking at these figures, we have that: (a) and (d) correspond respectively to the surface on the left and right side of the move with the simplest adapted handlebody structures; (b) is obtained from (a) by 1/2-handle addition, followed by 2-handle sliding only in Figure 35; (c) and (d) are obtained in turn by 2-handle slidings and 1/2-handle cancellation. The same figures also apply to the second case, after we replace by k's all the i's in the upper half and the j's in the lower half (except for the labels of the dotted line in the middle). The third case is trivial and we leave it to the reader.
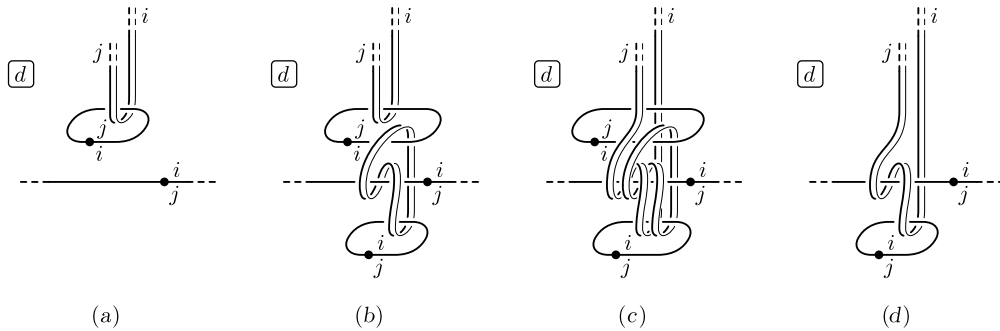

(a)   (b)   (c)   (d)
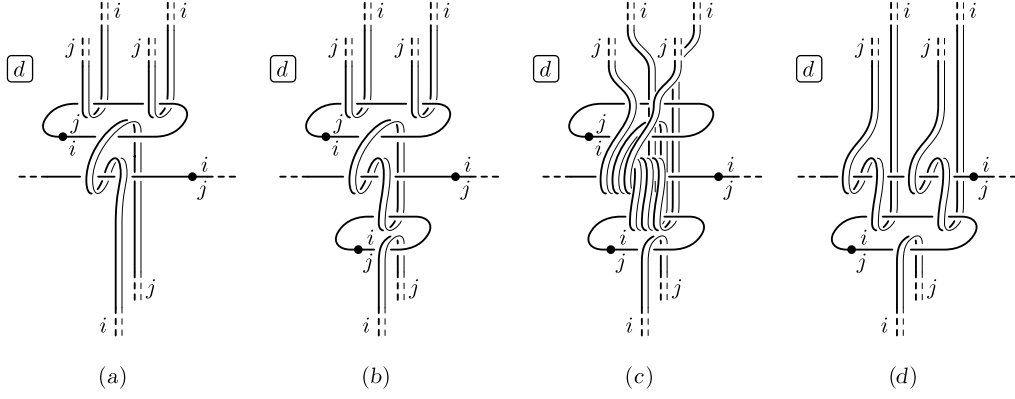
FIGURE 34.

FIGURE 35.

Finally, let us come to move I_4, which requires a bit more work than the other ones. As above, let $(i\ j)$ be the label of the horizontal band. Then, up to conjugation in Σ_d, there are eighteen possible ways to label the move, each one determined by the transpositions λ and ρ labelling respectively the left and right bottom ends of the diagonal bands. By direct inspection we see that, excluding the trivial cases when at least two of the three ribbon intersections involve bands with disjoint monodromies, which are left to the reader, and taking into account the symmetry of the move with respect to its inverse, there are only seven relevant cases: 1) $\lambda = (i\ j)$ and $\rho = (i\ j)$; 2) $\lambda = (i\ j)$ and $\rho = (i\ k)$; 3) $\lambda = (i\ k)$ and $\rho = (i\ j)$; 4) $\lambda = (i\ k)$ and $\rho = (i\ k)$; 5) $\lambda = (i\ k)$ and $\rho = (i\ l)$; 6) $\lambda = (i\ k)$ and $\rho = (j\ l)$; 7) $\lambda = (i\ k)$ and $\rho = (k\ l)$.

Figure 36 regards case 1. Here, (a) and (c) correspond respectively to the surfaces on the left side and right side of the move with suitable adapted handlebody structures, while (b) is related to (a) by two 2-handle slidings and to (c) by labelled isotopy. This figure also applies to case 4, after the same label replacement as above.
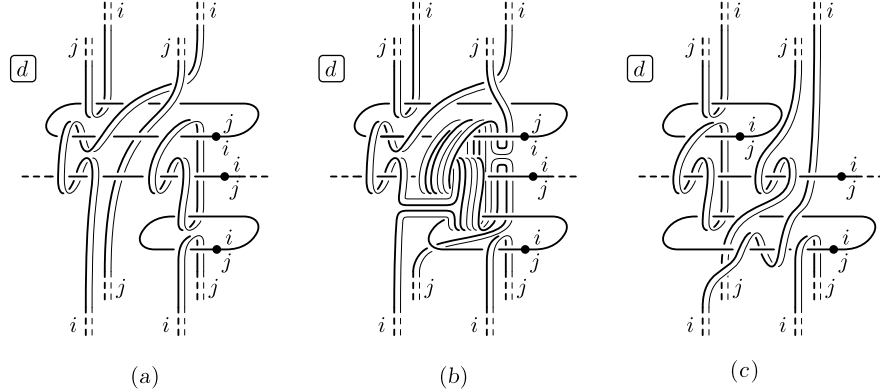

FIGURE 36.

Similarly, Figure 37 concerns with case 2 and, after the appropriate label replacements, also with cases 3 and 5. This time only one 2-handle sliding is needed to pass from (a) to (b). Figures 38 and 39 complete the proof, by dealing with the remaining cases 5 and 7. The three diagrams of Figure 38 are related by 1/2-handle addition/deletion, while the two diagrams of Figure 39 by labelled isotopy. $\square$
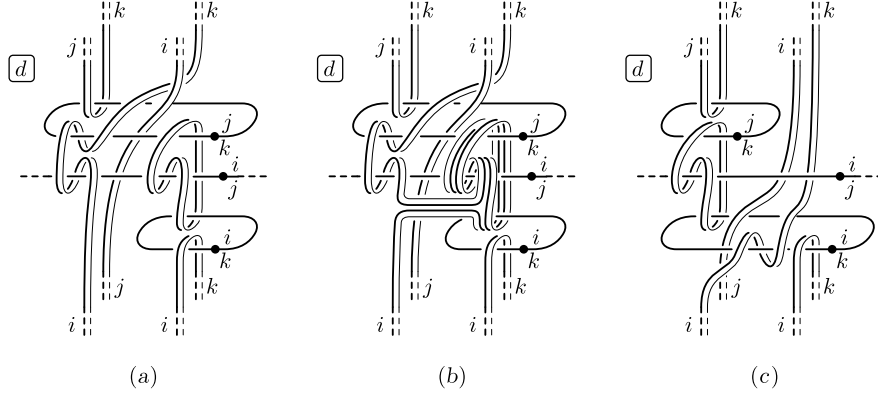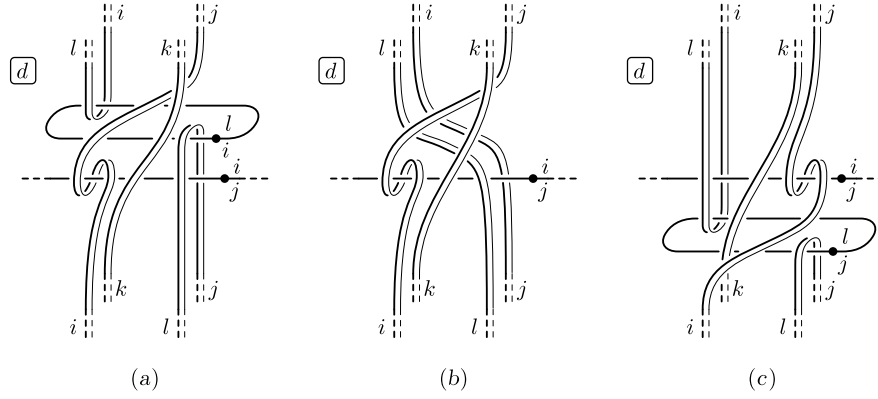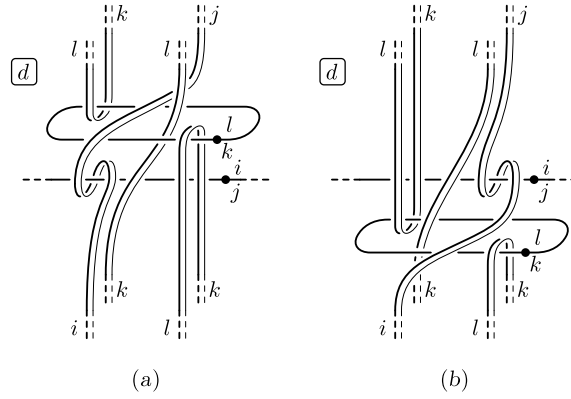
– 30 –

FIGURE 37.

FIGURE 38.

FIGURE 39.

It is worth remarking that Lemma 2.3 becomes trivial if we limit ourselves to require that the 4-dimensional 2-handlebodies represented by K_F and $K_{F'}$ are diffeomorphic, without insisting that they are 2-equivalent. In fact, labelled isotopy between F and F' (instead of labelled 1-isotopy) suffices for that, since it induces equivalence between the corresponding branched coverings, as recalled in Section 1. The relation between isotopy and 1-isotopy of ribbon surfaces in B^4 on one hand

and diffeomorphism and 2-equivalence of 4-dimensional 2-handlebodies on the other hand, will be discussed in Section 5.

PROPOSITION 2.4. *If the labelled ribbon surfaces $F, F' \subset B^4$, representing simple branched coverings of B^4, are related by labelled 1-isotopy, stabilization and moves R_1 and R_2, then the generalized Kirby diagrams K_F and $K_{F'}$ are equivalent up to 2-deformation moves.*

Proof. Labelled 1-isotopy has been already considered in the previous lemma. From the definitions it is apparent that stabilizing the branched coverings represented by a labelled ribbon surface F means adding a cancelling pair of 0/1-handles to K_F (cf. Figures 19 and 26).

Concerning moves R_1 and R_2, if F and F' differ by such a move, then by making the right choices in the construction of K_F and $K_{F'}$ we get the same result up to labelled isotopy. This is shown in Figure 40 (to be compared with Figure 2) for move R_1. The analogous and even easier case of move R_2 is left to the reader. $\square$
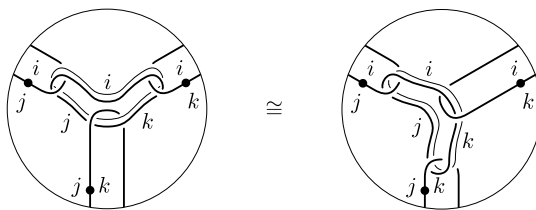

FIGURE 40.

We conclude this section with some further considerations on the ribbon moves R_1 and R_2. In particular, we see how they generate, up to labelled 1-isotopy, the auxiliary moves R_3, R_4, R_5 and R_6 described in Figure 41, where i, j and k are all distinct. These last moves will turn out to be useful in the next sections. First, we formalize in the following proposition the observation made in the Introduction about the inverses of moves R_1 and R_2, replacing isotopy by 1-isotopy.

PROPOSITION 2.5. *Moves R_1 and R_2 independently generate their own inverses, up to labelled 1-isotopy.*

Proof. For move R_1, the equation $R_1^{-1} = R_1^2$ obtained in the Introduction, by thinking R_1 as a rotation of 120°, holds also in the present context, since actually no isotopy is needed. On the other hand, move R_2^{-1} is equivalent, up to labelled diagram isotopy, to a suitable sequence of three moves of types I_2, I_3 and R_2 in the order. $\square$

PROPOSITION 2.6. *Moves R_1 and R_2 generate moves R_3, R_4, R_5 and R_6, as well as their inverses, up to labelled 1-isotopy.*

Proof. By Proposition 2.5 we do not need to worry about inverses. Move R_4 can be easily obtained as the composition of one move R_2^{-1} and one move R_2. Figures 42, 43 and 44 respectively shows how to get moves R_3, R_5 and R_6 in terms of labelled 1-isotopy and moves R_1. In Figure 42, we pass from (a) to (b) by one move I_2 and from (b) to (c) by one move R_1. In Figure 43, (b) is equivalent to (a) up to labelled diagram isotopy, then we perform respectively one move R_1, one move I_3 and one
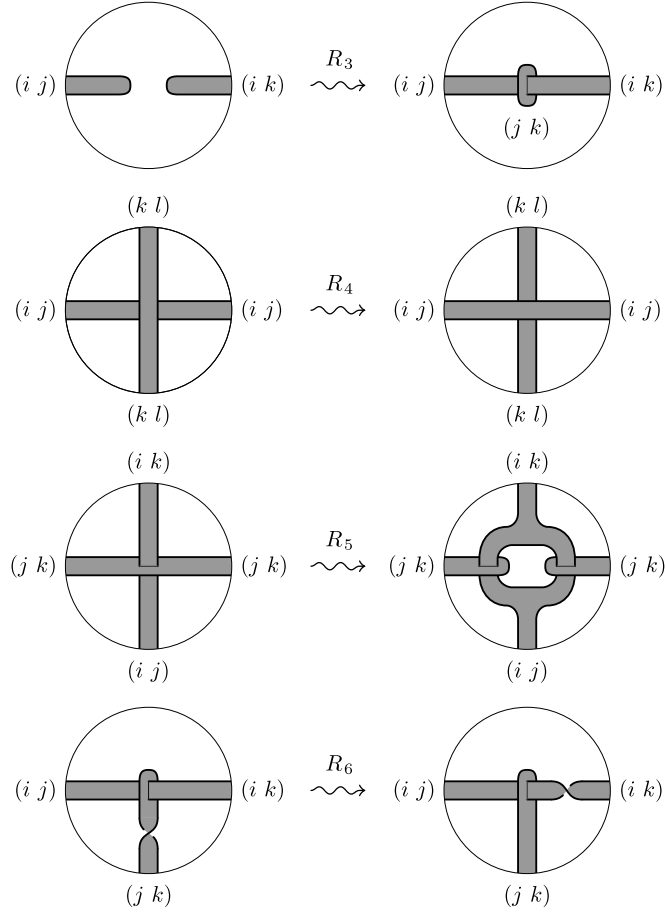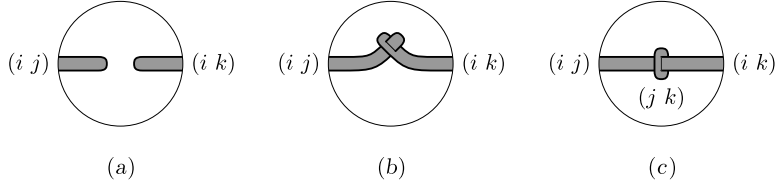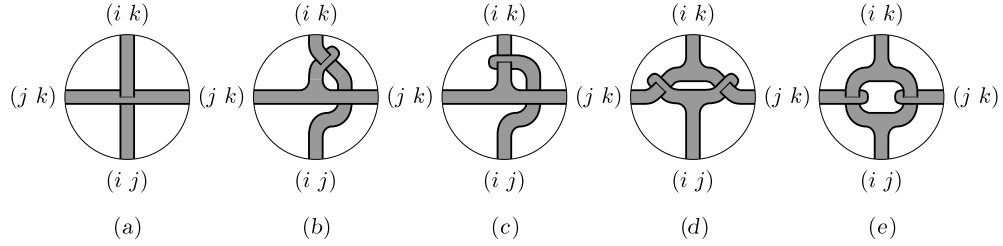
FIGURE 41.

FIGURE 42.

FIGURE 43.

pair of moves R_1 and R_1^{-1} to obtain in the order (c), (d) and (e). In Figure 44 we see that, up to conjugation by move R_1, the twist transfer of move R_6 can be realized by the labelled diagram isotopy between (b) and (c). $\square$
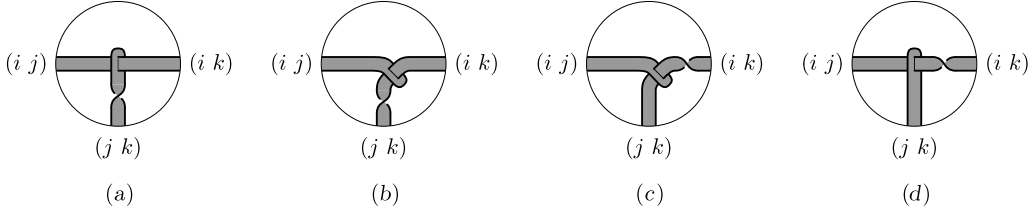
FIGURE 44.

REMARK 2.7. By labelled 1-isotopy and moves R_1 and R_2, any labelled ribbon surface representing a connected simple branched covering of B^4 can be made orientable, without changing the 2-equivalence class of the covering 4-dimensional 2-handlebody. In fact, twist transfer allows us to eliminate non-orientable bands as shown in Figure 45. Here, assuming $(i\ j)$ and $(j\ k)$ distinct, we pass from (a) to (b) by two moves of types I_2 and I_3, and from (b) to (c) by one move R_6^{-1}.
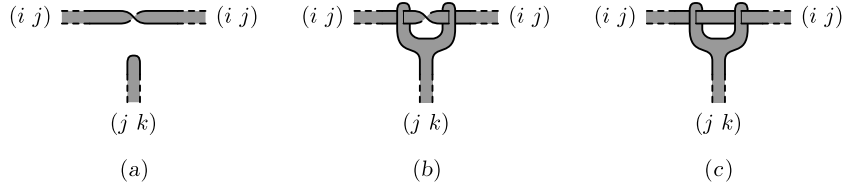

FIGURE 45.

3. From Kirby diagrams to labelled ribbon surfaces

In this section we prove the surjectivity of the map defined in the previous one, which associates to each simply labelled ribbon surface F the 2-equivalence class of the generalized Kirby diagrams K_F. Since everything can be done componentwise and any generalized Kirby diagram of a connected 4-dimensional 2-handlebody is 2-equivalent to an ordinary one (cf. Section 1), we will focus on ordinary Kirby diagrams and we will come back to the general case in the last Proposition 3.6.

Namely, for any ordinary Kirby diagram K, we construct a labelled ribbon surface F_K which represents the 2-equivalence class of the corresponding 4-dimensional 2-handlebody as a 3-fold simple branched covering of B^4 (cf. Proposition 3.5). Such a construction is canonical in the sense that the 4-stabilization of F_K is uniquely determined up to labelled 1-isotopy and covering moves R_1 and R_2. In this sense, F_K actually depends only on the 2-equivalence class of K.

In the light of the results of Section 4, we will relax the vertical triviality condition for the link L' in step (c) of our construction to just triviality (see Remark 4.4). However, we temporarily impose the vertical triviality condition, in order to be able to prove that F_K does not depend (in the above sense) on the choice of L', without resorting to the results of the next section.

Given an ordinary Kirby diagram K describing a 4-dimensional 2-handlebody $H^0 \cup H_1^1 \cup \ldots \cup H_m^1 \cup H_1^2 \cup \ldots \cup H_n^2$, we let $C_1, \ldots, C_m$ be the disjoint flat disks spanned by the unknots corresponding to the 1-handles and $L_1, \ldots, L_n$ be the framed

loops corresponding to the 2-handles. Moreover, we put $L = L_1 \cup \ldots \cup L_n$ and think of it indistinctly as a link or as a link diagram. Then, the construction of the labelled ribbon surface F_K is accomplished by the following steps:

(a) isotope K into a standard form (cf. Figure 23);

(b) add to K two standard disks $A_0, B_0 \subset R^3$ as shown in Figure 46; A_0 and the bottom part of B_0 (the one parallel to A_0 in the diagram) are supposed to lie in a horizontal plane such that K is entirely contained in the half space above it; the height of the rest of B_0 and the C_i's varies according to the height of the arcs of L passing through them (cf. step (c) below and Figure 51);
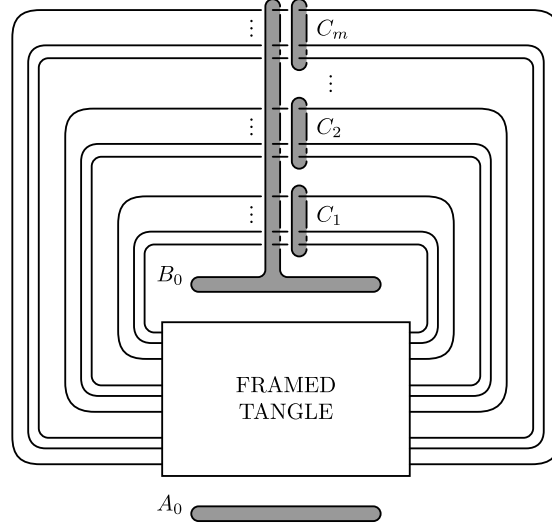

FIGURE 46.

(c) choose a vertically trivial state L' of L (cf. Remark 4.4) and call L'_i the component of L' corresponding to L_i; we think of L' as a vertically trivial link which coincides with L outside $E_1 \cup \ldots \cup E_l$, where each E_i is a cylinder projecting onto a small circular neighborhood of a changing crossing inside the tangle box of
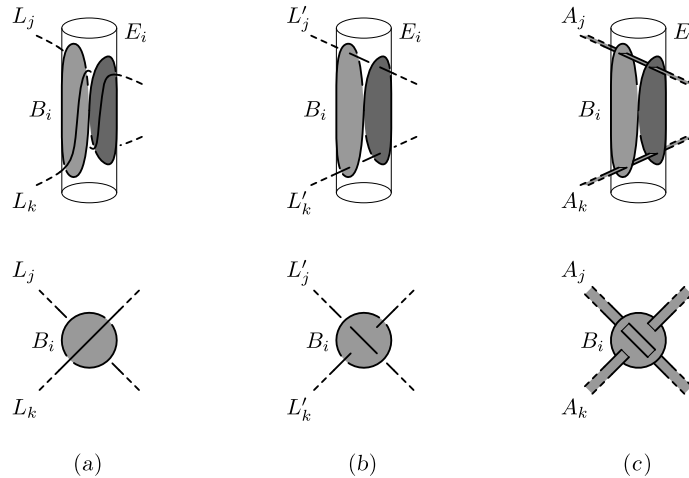

(a) (b) (c)

FIGURE 47.

Figure 46 (of course, this is possible only after having suitably vertically isotoped L); a cylinder E_i, together with the relative portion of diagram, is depicted in Figure 47 (a) and (b), where j and k may or may not be distinct; here $B_i \subset E_i$ is a regularly embedded disk without vertical tangencies, separating the two arcs of $L \cap E_i$ and forming four transversal intersection with L'; furthermore, we assume that the height function is strictly monotone and nearly constant on each arc of L' outside the tangle box and that it separates any two of such arcs, that is different arcs have disjoint height intervals;

(d) consider disjoint (possibly non orientable) narrow ribbons $A_1, \ldots, A_n \subset R^3$, such that each A_i has L'_i as its core and is obtained by a regular vertical homotopy from a ribbon representing half the framing of L_i plus one positive (resp. negative) full twist for each positive (resp. negative) crossing of L_i inverted to get L'_i; each A_i is assumed to be disjoint from A_0 and to form with the B_i's and the C_j's only the ribbon intersections shown in Figures 47 (c) and 48;
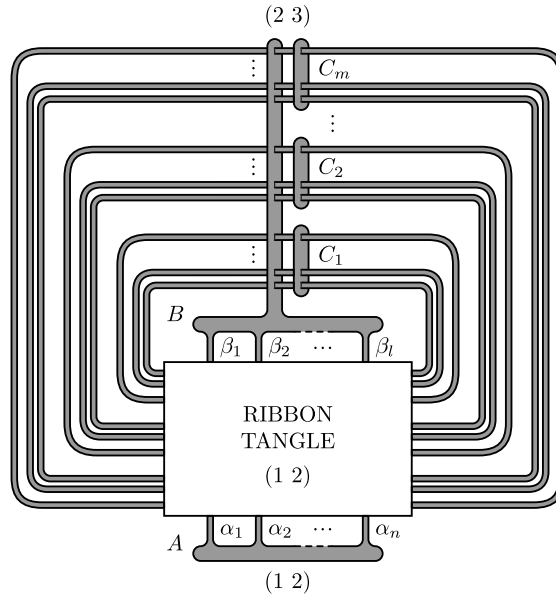

FIGURE 48.

(e) attach to $A_0 \cup A_1 \cup \ldots \cup A_n$ disjoint narrow bands $\alpha_1, \ldots, \alpha_n \subset R^3$ to get a connected non singular surface A; each α_i is an embedded 1-handle between A_0 and A_i, which is assumed to be disjoint from the E_j's and to have standard projection outside the tangle box as in Figure 48; we constrain the intersection of the α_i's with the tangle box to assume height values disjoint from the ones of the link L'; more precisely, if $[a_i, b_i]$ is the height interval of L'_i, then inside the tangle box α_i takes height values just below a_i; by the proof of Proposition 3.1, this last assumption is clearly much more that we really need, we nevertheless make it in order to keep things simpler;

(f) attach to $B_0 \cup B_1 \cup \ldots \cup B_l$ disjoint narrow bands $\beta_1, \ldots, \beta_l \subset R^3$ to get a connected non singular surface B; each β_i is an embedded 1-handle between B_0 and B_i, which is assumed to be disjoint from A and from the interiors of the E_j's and to have standard projection outside the tangle box as in Figure 48;

(g) define $F_K \subset B^4$ to be the ribbon surface having $A \cup B \cup C_1 \cup \ldots \cup C_m$ as 3-dimensional diagram, with the unique labelling assigning to A_0 the transposition (1 2) and to $B_0, C_1, \ldots, C_m$ the transposition (2 3) (cf. Figure 48); in particular, Figure 49 shows such labelling in a neighborhood of B_i (here, as in Figure 47, j and k may or may not be distict).
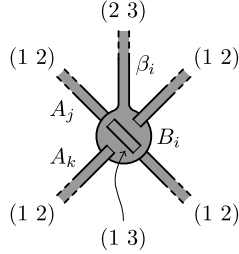

FIGURE 49.

It is worth noting that, by Remark 2.7 one more step could be added in order to make the ribbon surface F_K orientable and even more to make its diagram blackboard parallel. However, we omit such additional step, since we will not need those properties in what follows.

PROPOSITION 3.1. *Let K an ordinary Kirby diagram. Then the 4-stabilization of the labelled ribbon surface F_K constructed above is uniquely determined by K, that is it does not depend on the choices involved in the construction, up to labelled 1-isotopy and moves R_1 and R_2.*

Proof. First of all, as a preliminary, we add some extra structure to the above construction of F_K. Namely, we consider disjoint disks $D_1, \ldots, D_n \subset R^3$ respectively spanned by $L'_1, \ldots, L'_n$, such that the intersection of $D_1 \cup \ldots \cup D_n$ with any horizontal plane is either empty, one point or one arc. We constrain all these arcs to project onto the line segments drawn in Figure 50 outside the tangle box of Figure 46.
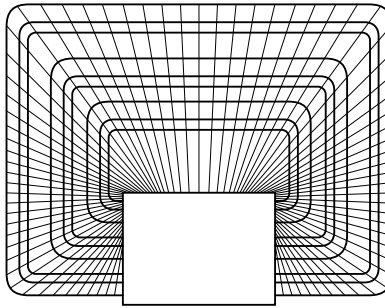

FIGURE 50.

Then, the portion of the D_i's outside the tangle box turns out to be nearly horizontal and completely determined by L'. Moveover, a suitable choice of the height function of B_0 and of the C_j's, ensures that the D_i's form with $B_0 \cup C_1 \cup \ldots \cup C_m$ only clasps and ribbon intersections, like the ones depicted in Figure 51, where the 3-dimensional view on the left side is compared with the corresponding diagram on the right side. Here, the clasps arise from the intersections of the L_i's with $B_0 \cup C_1 \cup \ldots \cup C_m$ and

the ribbon intersections are due to the height variations of B_0 and C_j's needed to let the L_i''s pass through them.
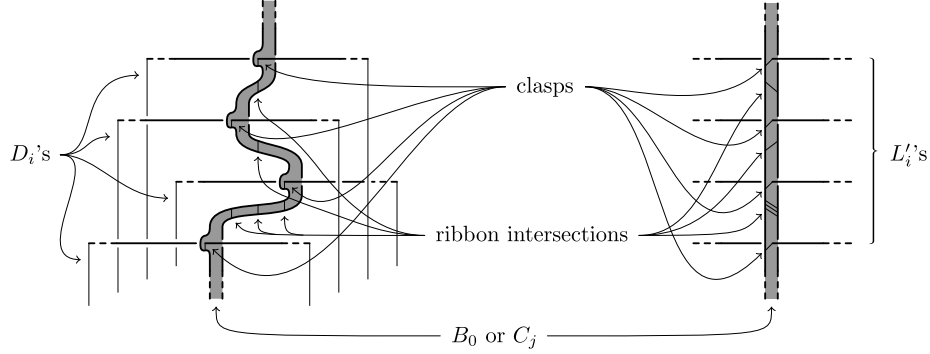
Without loss of generality, up to small perturbations, we can require that also inside the tangle box the D_i's form only clasps and ribbon intersections with F_K. Namely, we assume that: 1) $A_j \cap D_j$ consists of a certain number of disjoint clasps connecting L_j' with the boundary of A_j, in such a way that $D_j \cup A_j$ is collapsible; 2) each B_j forms with the D_i's four clasps and some (possibly none) ribbon intersections, as shown in Figure 52; 3) the β_j's may pass through the D_i's, forming ribbon intersections with them. Finally, we observe that, by construction, each D_i is disjoint from the A_j's with $j \neq i$ and from all the α_j's.
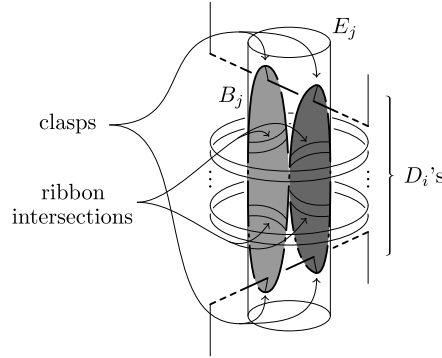

FIGURE 52.

At this point we pass to the core of the proof. Given an ordinary Kirby diagram K, the relevant choices occurring in the construction of F_K are, in order: 1) the standard form K; 2) the vertically trivial state L'; 3) the bands $\alpha_1, \ldots, \alpha_n$; 4) the bands $\beta_1, \ldots, \beta_l$. In fact, the A_i's and the B_i's are uniquely determined up to diagram isotopy.

We prove that the 4-stabilization of F_K is independent on these choices, up to labelled 1-isotopy and moves R_1 and R_2, by proceeding in the reverse order and assuming each time that all previous choices have been fixed. By Propositions 2.5 and 2.6, in addition to moves R_1 and R_2, we can use also the moves R_3, R_4, R_5 and R_6 introduced in the previous section, as well as the inverses of all such moves.

Concerning the α_i's and the β_i's, it suffices to prove that, in presence of a stabilizing disk, labelled 1-isotopy and the moves above enable us to change them one by one.

Figure 53 shows how to deal with the band β_i of Figure 49. The small disk with label (3 4) in (a) is the stabilizing disk. This can be moved to form one ribbon intersection with β_i as in (d), by labelled 1-isotopy and four moves R_2. Parts (b) and (c) of the figure represent 1-isotopic intermediate steps. Looking at the diagram, we can realize such a modification by an isotopy $H : B^2 \times [0,1] \to R^3$ between the two disks labelled (3 4) in (a) and (d), which is a suitable homeomorphism of $B^2 \times [0,1]$ onto a regular neighborhood of B_i whose boundary forms four ribbon intersections with A_i and A_j. Finally, we perform one move R_3^{-1} on (d) to cut the band β_i. The result is clearly independent of β_i up to diagram isotopy, so we are done.
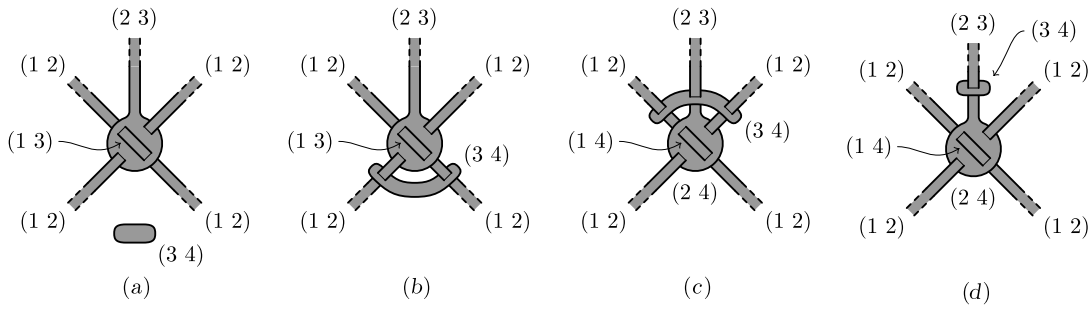
FIGURE 53.

The same idea also applies to the band α_i, as suggested by Figure 54. We warn the reader that Figure 54 is much more sketchy than Figure 53. In fact, in place of the disk B_i we have here the complex $D_i \cup A_i$, which can be large and complicated, although still collapsible. Before of starting the process, we let the stabilizing disk of Figure 53 (a) pass first through B_0 and then through A_0, in such a way that its label becomes (1 4), as in Figure 54 (a). This time, the need for move R_2 is due to

FIGURE 54.

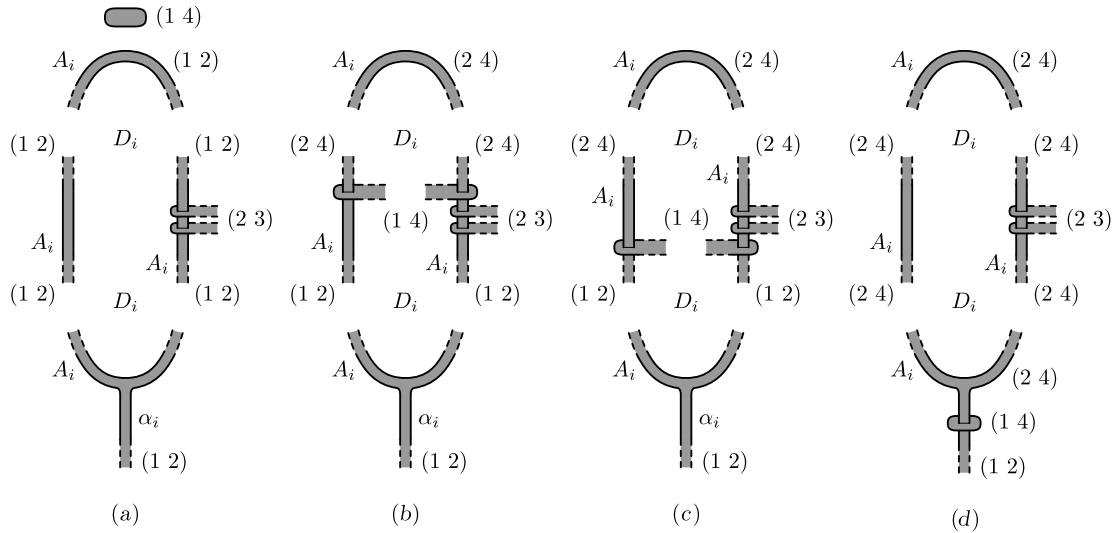the ribbon intersections that A_i may form passing through the B_j's and the C_j's and that the B_j's (including B_0), the C_j's and the β_j's may form passing through the interior of D_i (cf. Figures 51 and 52). In particular, the ribbon intersections along A_i always appear in pairs, each pair being formed with B_0 and one of the C_j's (cf. Figure 48) or with one of the B_j's (cf. Figure 47). Any such pair looks like the one pictured in Figure 54 (a). Comparing steps (b) and (c), one see how the stabilizing disk can be pushed beyond this pair, by using labelled 1-isotopy and moves R_2. On the other hand, only one move R_4 suffices to go beyond each one of the ribbon intersections in the interior of D_i. Eventually, the stabilizing disk reach the position of step (d), so we can conclude as above by cutting the band α_i.

Now we pass on to the vertically trivial state L'. Recall that we are thinking of it as a vertically trivial link, that is a vertically trivial diagram together with a compatible height funtion. Of course, course different choices of the height function compatible with the same diagram are related by a vertical diagram isotopy. Such an ambient isotopy can be used to relate the entire resulting surfaces except for the bands $\alpha_1, \ldots, \alpha_n$. In fact, each band α_i is forced, by the vertical constraints we imposed in its definition (cf. step (e) of the construction of F_K), to be attached to the corresponding A_i near to the point of minimum height. However, this problem can be overcome by the above proof of independence on the α_i's. Actually, this delicate point is the only obstruction to the naive solution of the problem of the independence on the α_i's consisting in fixing a standard form for them.

Thus, having settled the problem of the height function, we are left with the modifications of Proposition 1.1 on the diagram. First we address the change of ordering of the link components. Of course, it is enough to deal with the transposition in the vertical order of any two components L'_i and L'_j, that means to simultaneously change all the crossing involving both L'_i and L'_j. Without loss of generality we assume that L'_i lies under L'_j. We begin as above, operating with the stabilizing disk around the D_i to get the configuration of Figure 54 (d). As a result, we have a global labelling change on the ribbon A_i from (1 2) to (2 4), while the labelling of A_j is left unchanged. Then, we can perform the crossing changes as described in Figure 55. Here, apart from 1-isotopy, we have only one move R_4 relating (b) and (c). To be precise, Figure 55 covers only one of the two possible cases, the other one being covered by the same steps in the reverse order with the roles of A_i and A_j exchanged. After all the crossing changes have been performed, we bring the stabilizing disk back to the original position, by reversing the process of Figure 54.
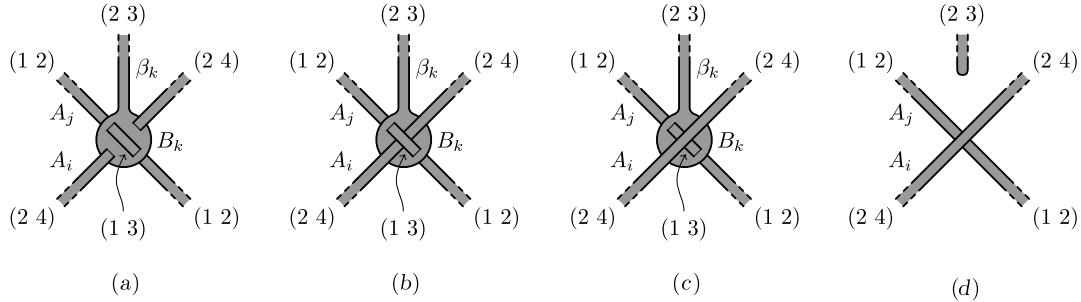

FIGURE 55.

Concerning single crossing changes making a vertically trivial component L'_i into a different vertically trivial state of L_i, there are four cases to be considered, depending on sign of the crossing and on whether L'_i coincides with L_i at that crossing or not. In all cases, it is not restrictive to assume that the two points of L'_i projecting to the crossing are distinct from the unique minimum height point p_i of L'_i and that the vertical segment joining them is contained in D_i. Such a segment divide D_i into two disks. We call D'_i the one which does not contain p_i and assume that the attaching arc of the band α_i to A_i is disjoint from the part of A_i running along the boundary of D'_i.

Figure 56 indicates how to realize the crossing change in one of the four cases. For the other three cases it suffices to apply a mirror symmetry to all the stages and/or reverse their order. First, we pass from (a) to (b) by the same the process described in Figure 54, with the only difference that here we have D'_i in place of D_i. We continue that process one more step to get (c), by pushing the stabilizing disk beyond one of the two pairs of ribbon intersections between A_i and B_j. Then, we obtain (d) from (c) in the same way we obtained (c) from (a) in Figure 55. Finally, we push back the stabilizing disk through the same ribbon intersections as above to achieve (e). At this point, we bring back the stabilizing disk in the original position, by reversing the process from (a) to (b), after having transferred the full twist present on it to A_i by move R_6. We remind the reader that the additional full twist on A_i compensates the change of crossing (cf. definition of A_i in step (d) of the construction of F_K).
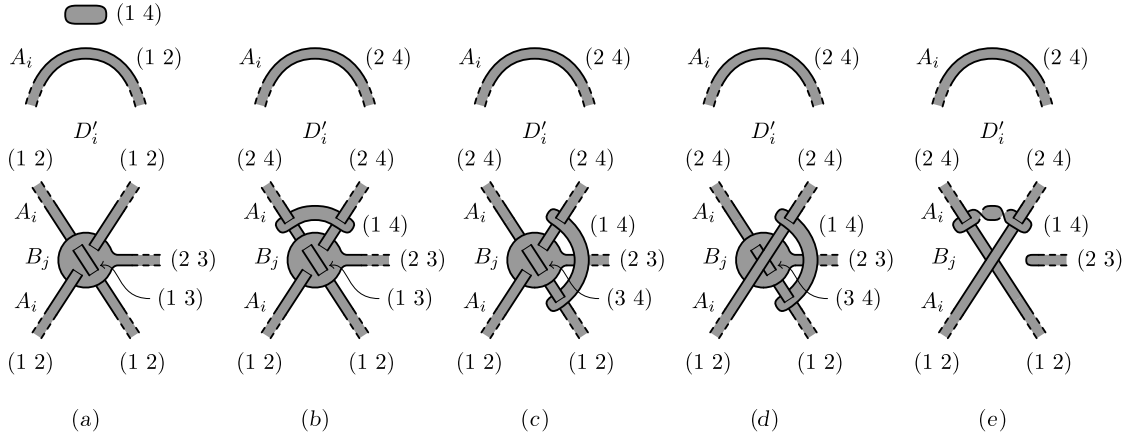

FIGURE 56.

It remains to deal with the standard form of K. As discussed in Section 1, different choices of such standard form are related by a finite sequence of Reidemaister moves inside the tangle box and moves of the types depicted in Figures 24 and 25. Hence, we have to prove the invariance of the 4-stabilization of F_K with respect to those moves.

We observe that any Reidemeister move on L induces the same move on L' and so just a diagram isotopy on F_K, provided that none of the involved crossings (before as well as after the move) has been changed when passing from L to L'. The reason is that in this case the two links coincide inside a small 3-cell where the move takes place and such 3-cell is free from the B_i's. We leave to the reader the

straightforward verification that a vertically trivial state L' of L with the required property can be always achieved by a suitable application of the naive unknotting procedure described in Section 1 (with height function on each component as in Figure 8 (a) or (c), depending on the move).

On the other hand, by relaxing a little bit the constraints for the standard form of K outside the tangle box, the moves of Figures 24 and 25 reduce to the ones given in Figure 57.
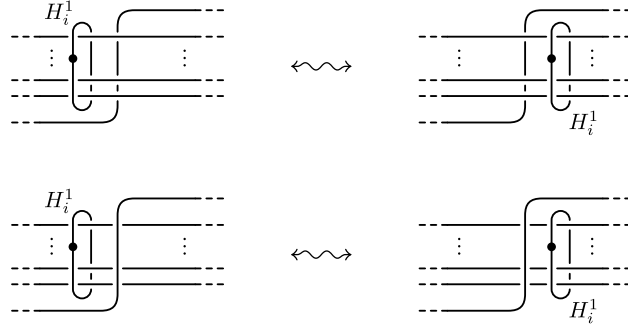
FIGURE 57.

To be more precise, let us look at the following Figures 58 and 59, which regard respectively the moves of Figures 24 and 25. Here, only the relevant part of the diagram is drawn, however it should be clear what is happening. We notice that, in addition to Reidemeister moves inside the tangle box, we just need to bring some crossings out of the tangle box by isotopy of the planar link diagram and to move certain dotted components through them by a sequence of moves as in Figure 57

FIGURE 58.

FIGURE 59.

(for the sake of simplicity, the figures are drawn as if the Reidemeister moves were performed directly outside the tangle box).

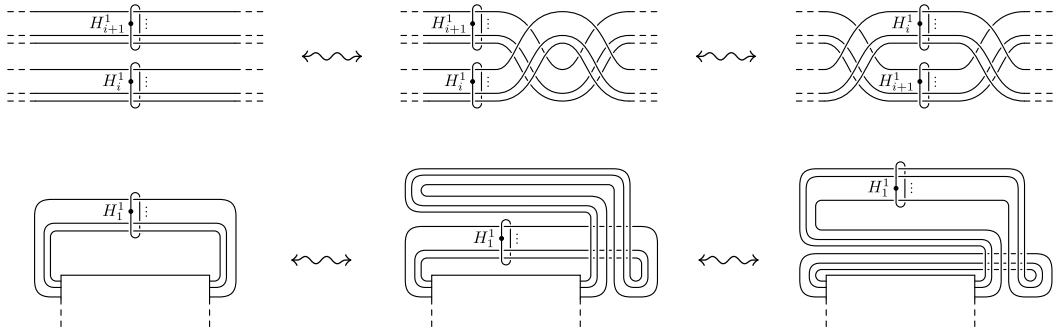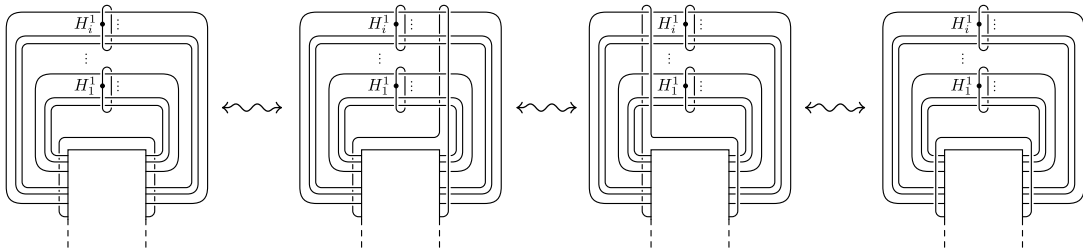To conclude the proof, we have only to interpret the moves of Figure 57 on the diagram K in terms of moves R_i on the corresponding surface F_K. In fact, we already know how to handle Reidemeister moves, while isotopy of the planar link diagram of K obviously induces diagram isotopy on F_K.

We first operate on the disks $C_i, \ldots, C_{i+h}$ corresponding to the 1-handles we want to move, as described in Figure 60. Namely, apart from the diagram isotopy relating (a) and (b), we perform a certain number of moves R_2 to let the stabilizing disk labelled by $(3\,4)$ reach the position of (c) and then we create the new component B_0' with monodromy $(2\,4)$ that appears in (d) by one move R_3. We emphasize that also the monodromy of the disks $C_i, \ldots, C_{i+h}$ is now $(2\,4)$. Actually, we could limit ourselves to one single disk at a time, but for our purpose it is more convenient to operate simultaneously on all the disks involved by the moves of Figures 58 and 59.
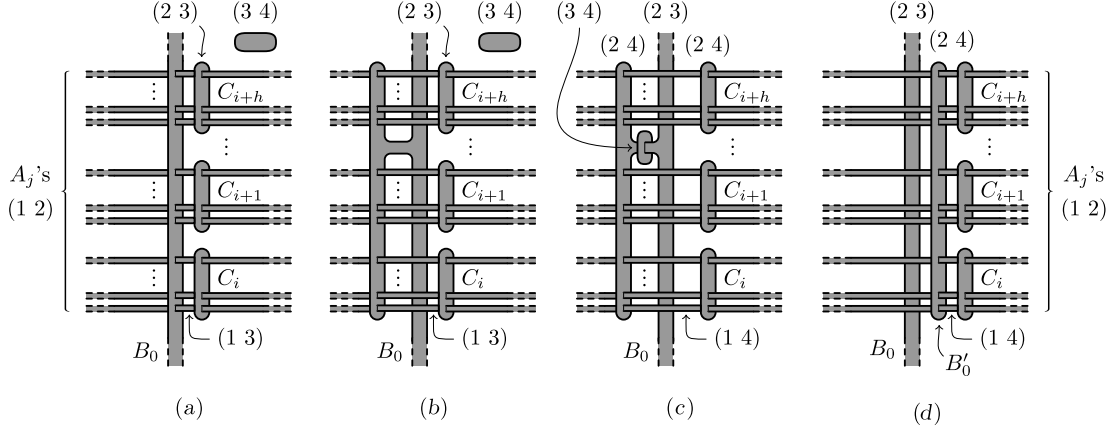
FIGURE 60.

Then, we observe that the moves of Figure 57 would correspond to a diagram isotopy of F_K, if L' coincided with L at all the involved link crossings. In fact, if this were the case it would be enough to slide B_0' and C_i over/under all such crossings. So, we only have problems with the disks B_k at the crossings we changed to get the vertically trivial link L'. Figure 61 shows how to handle these crossings. Here, apart from 1-isotopy, we apply four moves R_4 to pass from (b) to (c). □

FIGURE 61.

REMARK 3.2. The last part of the proof of Proposition 3.1 suggests that the above construction of the labelled ribbon surface F_K could be easily adapted to work directly with any ordinary Kirby diagram, avoiding the neeed for the standard form. A little more effort would enable us to include also generalized Kirby diagrams, of course by using labellings of arbitrary degree (cf. proof of Proposition 3.6). The reason for fo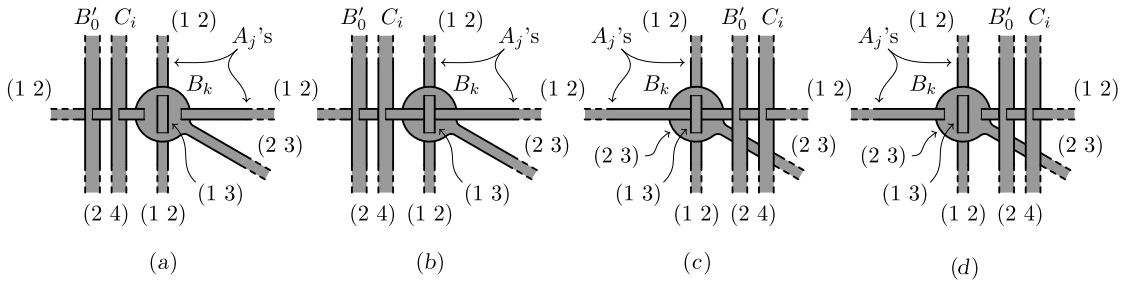cusing on the ordinary case is that we want to control the degree of the converings, in view of the results of the next section. On the contrary, the choice of working with the standard form is motivated only by the sake of convenience.

So far we have proved that the labelled ribbon surface F_K associated to a Kirby diagram K, is well defined up to 4-stabilization, labelled 1-isotopy and ribbon moves. The next proposition addresses the invariance of F_K under 2-deformations of K.

PROPOSITION 3.3. *If K and K' are 2-equivalent ordinary Kirby diagrams, then the 4-stabilizations of the labelled ribbon surfaces F_K and $F_{K'}$ are equivalent up to labelled 1-isotopy and moves R_1 and R_2.*

Proof. By Proposition 3.1 we can assume K and K' to be in standard form and the 2-deformation between them to be described in terms of moves as in Proposition 1.4. Moreover, we do not need to worry about handle isotopy, since diagram isotopy has already been treated in the proof of Proposition 3.1, while the first move of Figure 18 performed on ordinary Kirby diagrams in standard form trivially induces labelled isotopy on the corresponding labelled ribbon surfaces.

Hence, it remains to consider the cases when K and K' differ by a pair of cancelling 1/2-handles (cf. Figure 19) and by a 2-handle sliding (cf. Figure 20).

The first case is quite easy. In fact, if the disk C_i and the loop L_j of K represent two cancelling handles H_i^1 and H_j^2, then in F_K only the ribbon A_j passes once through C_i. Therefore, by a move R_3 we can remove C_i and break A_j into two long tounges. At this point, labelled 1-isotopy allows us to completely retract such tounges and after that also the ones $\beta_k \cup B_k$ related to crossings involving L_j.

The case of a 2-handle sliding requires some preliminaries. First of all, let us renumber the 2-handles starting from the two ones involved in the sliding, in such a way that H_1^2 slides over H_2^2. In terms of Kirby diagram, this means to replace L_1 with the band connected sum $L_1 \#_\gamma \bar{L}_2$, where $\bar{L}_2$ is a parallel copy of L_2 realizing its framing, and γ is a band connecting L_1 to $\bar{L}_2$. Up to isotopy, γ can be assumed to be a blackboard parallel band which does not form any crossing with the L_i's, as in Figure 62 (a). We also assume that the vertically trivial status L' choosen to construct F_K satisfies the following properties: 1) the vertical order of the components is the one given the numbering, that is L_i' lies under L_j' for any $i < j$; 2) the minimum point p_1 (resp. p_2) and the maximum point q_1 (resp. q_2) of the
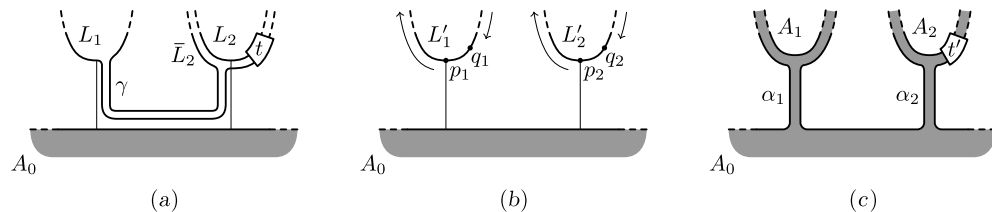

FIGURE 62.

– 44 –

height function on L_1' (resp. L_2') coincide with the end points of (resp. are close to) the attaching arc of γ to L_1 (resp. $\overline{L}_2$), as in Figure 62 (b). Here, the arrows indicate the orientations that we will use in the framing computation at the end of the proof, so they are not relevant for the moment. Finally, we choose α_1 and α_2 to be blackboard parallel bands, such that γ can be thought to run parallel to them and to the part of Bd A_0 between them, as in Figure 62 (c). For the sake of convenience, the framing of L_2 and the ribbon A_2 are assumed to be blackboard parallel outside the twist boxes t and t' respectively in Figure 62 (a) and (c). We warn the reader that the number of twists inside such boxes is not the same in (a) and (c), accordingly to step (d) of the construction on F_K.

Once it has been set up in this way, the sliding can be interpreted in terms of ribbon moves on the 4-stabilization of F_K as sketched in Figure 63. We think of A_1 as a 1-handle attached to $\mathrm{Cl}(A - A_1)$ and slide one of its attaching arcs along the boundary of $\mathrm{Cl}(A - A_1)$ as indicated by the arrows in (a) and (b). Before of reaching the twist box t, this sliding can be entirely realized by labelled diagram isotopy, except for the labelled 1-isotopy moves (of types I_2 and I_3) needed to pass through the disks B_i encountered by A_2. Each time a disk B_i is passed through, two new ribbon intersections appear as shown in part (a) of Figure 64. Then, we use again 1-isotopy to split B_i into two twin disks similar to the original one, as suggested by the remaining parts of Figure 64.
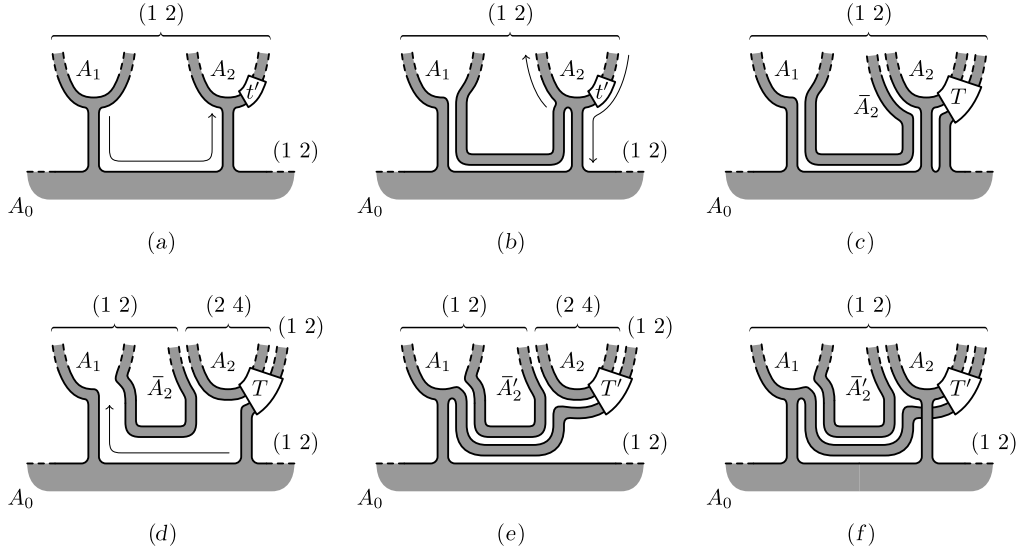
FIGURE 63.

When traversing the twist box t' in (b) to get the twist box T in (c), after having followed all the twists of A_2, we add some further crossings between A_2 and the parallel closed ribbon $\overline{A}_2$ (together with further B_i's), in order to make them unlinked. Actually, $\overline{A}_2$ itself is always well defined thanks to these additional crossings, having their number the same parity of the number of half twists in t'. Figure 65 shows how to add a positive crossing; for a negative one it suffices to mirror the figure. Here, some moves other than 1-isotopy are needed: one move R_5 from (a) to (b); two opposite moves R_6 (twist transfers) from (c) to (d); two moves R_1 from (d) to (e).

FIGURE 64.

FIGURE 65.

Then, we consider a disk D_2 spanned by L_2' as in the proof of Proposition 3.1 and perturb it near to A_2 in such a way that it becomes disjoint from $\overline{A}_2$, while remaining disjoint from all the oth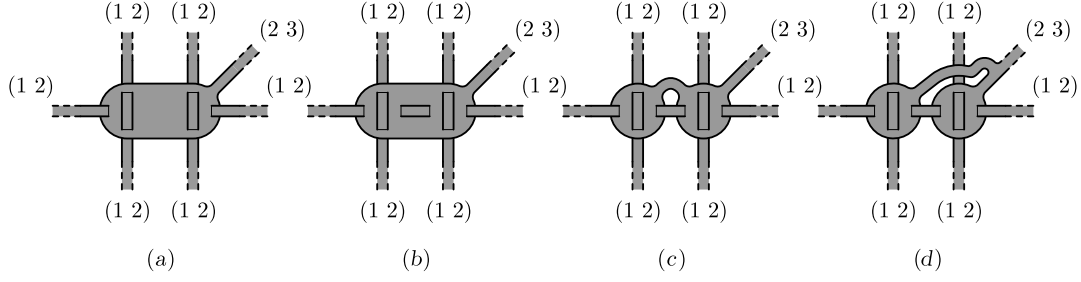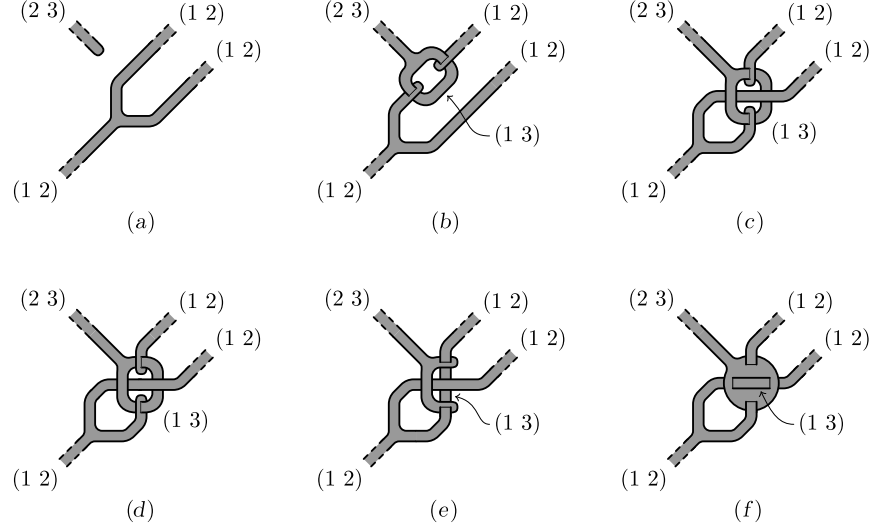er A_i's and continuing to form only clasps and ribbon intersections with the rest of the surface. Such a perturbed disk can be used in place of the original one in the process of Figure 54, to bring the stabilizing disk around α_2. After that, we cut α_2 by a move R_3^{-1} to obtain the labelled surface of Figure 63 (d). At this point, we can continue the sliding as indicated by the arrow and we can change all the crossings where $\overline{A}_2$ passes over A_2, by operating as in Figure 55. In this way we get (e), where $\overline{A}_2'$ and T' differ from $\overline{A}_2$ and T only by the performed crossing changes. To end up with (f), we first observe that $\overline{A}_2'$ crosses always under A_2, so it can be pushed down below the plane $z = a_2$ (recall that $[a_2, b_2]$ is the eight interval of L_2'). Hence, after having restored the band α_2 by a move R_3, we can use the process of Figure 54 in the opposite direction, this time with the original disk D_2, to take back the stabilizing disk.

Finally, we want to verify that the labelled ribbon surface resulting from all the above modifications coincides with $F_{K'}$, where K' is the ordinary Kirby diagram obtained from K by replacing L_1 with $L_1 \#_\gamma \overline{L}_2$.

Looking at Figure 63 (f), we call $\overline{L}_2'$ the core of $\overline{A}_2'$ and observe that here the original ribbon A_1 has been replaced by the ribbon $A_1 \#_\gamma \overline{A}_2'$ with core $L_1' \#_\gamma \overline{L}_2'$.

Taking into account the choices made at the beginning about the height function of L' and taking the care of preserving the vertical triviality of $\overline{L}'_2$ when pushing down $\overline{A}'_2$, we can assume that $L'_1 \#_\gamma \overline{L}'_2, L'_2, \ldots, L'_n$ form a vertically trivial link.

We claim that, up to isotopy, this is a vertically trivial status of the link formed by $L_1 \#_\gamma \overline{L}_2, L_2, \ldots, L_n$ and that the crossings at which the two links differ are exactly the ones marked by the presence of a disk B_i. In fact, it is clear from the construction that, by inverting such crossings in the vertically trivial link formed by $L'_1 \#_\gamma \overline{L}'_2, L'_2, \ldots, L'_n$, we get a link of components $L_1 \#_\gamma \widehat{L}_2, L_2, \ldots, L_n$, where $\widehat{L}_2$ is a certain parallel copy of L_2. Then, our claim reduces to asserting that $\overline{L}_2$ and $\widehat{L}_2$ represent the same framing of L_2, that is $\mathrm{Lk}(L_2, \overline{L}_2) = \mathrm{Lk}(L_2, \widehat{L}_2)$. This equality between linking numbers follows from some easy computations involving the writhes $w_2 = \mathrm{Wr}(L_2)$ and $w'_2 = \mathrm{Wr}(L'_2)$ and the signed number $c_2 = (w_2 - w'_2)/2$ of the crossings of L_2 inverted to get L'_2. Denoting by $f_2 = \mathrm{Lk}(L_2, \overline{L}_2)$ the framing of L_2 in K, we have $f_2 - w_2$ full twists inside the twist box t of Figure 62 and $f_2 + 2c_2 - 2w_2$ half twists inside the twist box t' of Figure 63 (see step (d) in the definition of F_K). As a consequence, the additional crossings we inserted inside the twist box T of Figure 63 is $-2w'_2 - (f_2 + 2c_2 - 2w_2) = 2c_2 - f_2$. Then, the signed number of the crossings between A_2 and $\overline{A}_2$ marked by the B_i's is $-f_2$, being $-2c_2$ the signed number of such crossings outside the twist box T. Since both $\overline{A}_2$ and $\overline{A}'_2$ are unlinked from A_2, this number of crossings remains unchanged if we replace $\overline{A}_2$ with $\overline{A}'_2$ and we can conclude that $\mathrm{Lk}(L_2, \widehat{L}_2) = f_2$.

It remains to check that the ribbon $A_1 \#_\gamma \overline{A}'_2$ in Figure 63 (f) represents the right half integer framing of $L'_1 \#_\gamma \overline{L}'_2$. To do that, we orient L_1, L_2, L'_1 and L'_2 accordingly to Figure 62 (b). Then, the signed number of crossings of $L_1 \#_\gamma \overline{L}_2$ to be inverted in order to get $L'_1 \#_\gamma \overline{L}'_2$ is $c_1 + c_2 + \mathrm{Lk}(L_1, L_2)$, where c_1 is defined analogously to c_2. On the other hand, the framing of $L_1 \#_\gamma \overline{L}_2$ in K' is $f_1 + f_2 + 2\mathrm{Lk}(L_1, L_2)$, where f_1 is the framing of L_1 in K, and so $A_1 \#_\gamma \overline{A}'_2$ should be equivalent up to vertical regular homotopy to a ribbon representing the half integer framing $f_1/2 + f_2/2 + c_1 + c_2 + 2\mathrm{Lk}(L_1, L_2)$. The reader can easily realize that this is the case, taking into account that $\mathrm{Wr}(L_1 \#_\gamma \overline{L}_2) = w_1 + w_2 + 2\mathrm{Lk}(L_1, L_2)$, where $w_1 = \mathrm{Wr}(L_1)$. $\square$

REMARK 3.4. Let us recall that any crossing in a Kirby diagram K can be inverted, up to 2-deformation, by adding a suitable pair of 1/2-handles, as shown in Figure 66. In the light of the preceding proposition, any disk B_i in the 4-stabilization of F_K can interpreted, up to labelled 1-isotopy and ribbon moves, as such a pair of 1/2-handles. A direct proof of this fact is provided in Figure 67. Here, apart from labelled 1-isotopy, we perform two twist transfers (moves $R_6^\pm$) to get (c) from (b) and one move R_5 followed by one more twist transfer to get (d) from (c).
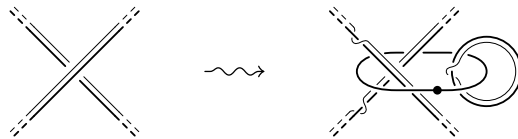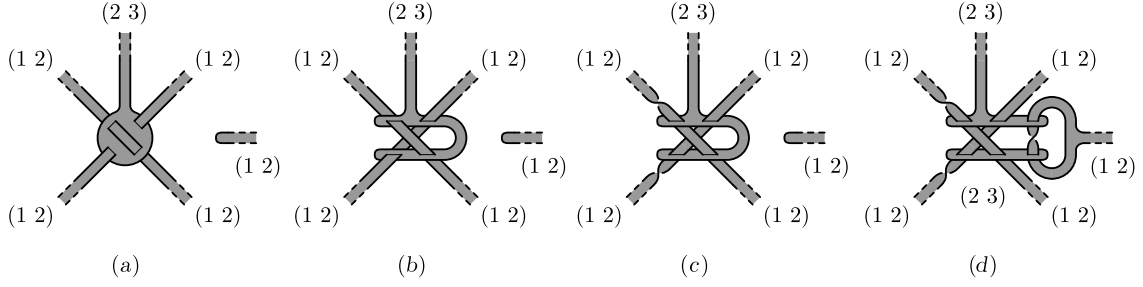

FIGURE 66.

FIGURE 67.

To complete this section we are left with showing that, up to 2-deformations, the construction of F_K given above is inverted by that one of K_F given in the previous section. As observed at the beginning of the section, this implies that any 2-equivalence class of 4-dimensional 2-handlebodies can be represented as a simple branched covering of B^4 (cf. Proposition 3.6). In particular, we can insist that the covering has degree 3 in the case of connected handlebodies.

PROPOSITION 3.5. *Let K be an ordinary Kirby diagram and $F = F_K$ be the corresponding labelled ribbon surface. Then, the generalized Kirby diagram K_F is equivalent to K up to 2-deformation moves.*

Proof. Recall that for constructing K_F one need first to choose an adapted 1-handlebody structure on F, even if the 2-equivalence class of K_F is independent on this choice by Proposition 2.2.

We claim that there exists an adapted 1-handlebody structure on F, naturally related to the above construction, such that the generalized Kirby diagram K_F constructed starting from it is equivalent to K up to labelled isotopy, 1-handle slidings and deletion of cancelling 0/1-handles.

Without loss of generality, we suppose that K is already in the standard form choosen in step (a) of the construction of F_K. Moreover, we adopt all the notations introduced during that construction.

To specify the claimed adapted handlebody structure of F, we first decompose each ribbon $A_i \subset F$ as $A_i^0 \cup A_i^1$, where A_i^0 is a small 0-handle containing the attaching arc of α_i and A_i^1 is a 1-handle. Then, we consider the adapted 1-handlebody structure of F whose 0-handles are $A_0 \cup \alpha_1 \cup \ldots \cup \alpha_n \cup A_1^0 \cup \ldots \cup A_n^0$, B, $C_1, \ldots, C_m$ and whose 1-handles are $A_1^1, \ldots, A_n^1$ (cf. Figure 48). By a suitable choice of the α_i's and the β_i's we can assume that all the 0-handles are blackboard parallel.

The generalized Kirby diagram K_F constructed starting from this handlebody structure is sketched in Figure 68. Here, as well as in all the figures of this proof, we omit to draw the framings for the sake of readability. Some further details of K_F are shown in Figures 69 and 70. The labelled isotopy modifications described there are performed at all the α_i's and at all the crossings between the A_i's.

Once such modifications have been performed, we slide all the 1-handles corresponding to the C_i's over the one corresponding to B, in such a way that, up to diagram isotopy, we are left with the diagram of Figure 71. Here we have two overlapping but vertically separated tangle boxes T (in front) and T' (in back), respectively labelled by 2 and 1.
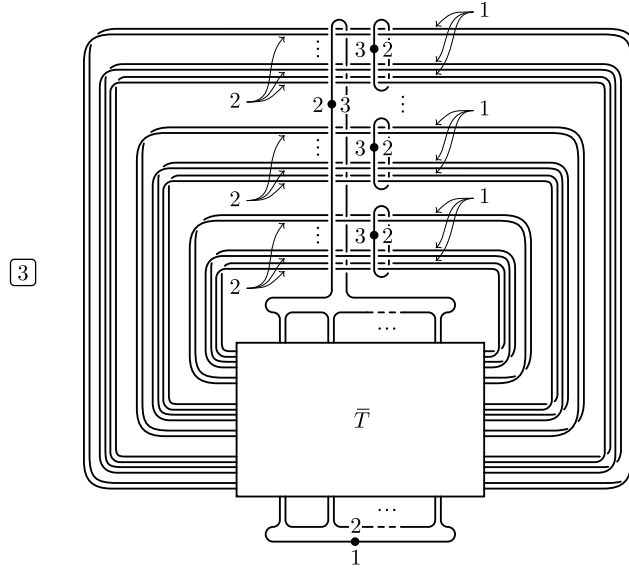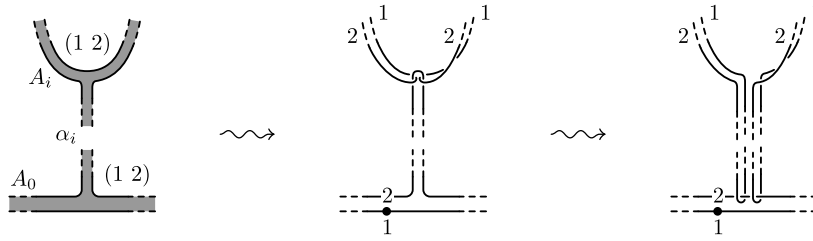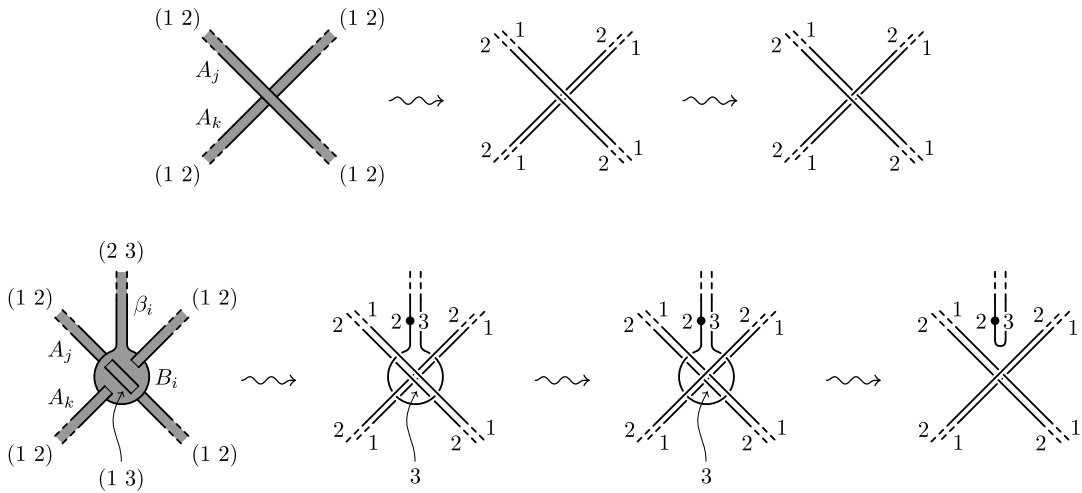
FIGURE 68.

FIGURE 69.

FIGURE 70.

Disregarding for the moment the framings, the link formed by the undotted components in Figure 71 is the componentwise band connected sum of the original link L and a parallel copy L'' of its vertically trivial state L' pushed down to cross under everything else (including the dotted components). Each component L_i of L

is connected to the corresponding component L_i'' of L'' by a band γ_i running back and forth on the two sides of α_i. Since L'' is trivial and unlinked from the rest of the diagram and the bands γ_i can be assumed to be disjoint from a set of trivializing disks for L'', we can isotope the diagram to get back K entirely labelled by 2 with two extra dotted components labelled by 1,2 and 2,3 separated from it. Finally, such dotted components can be eliminated by 0/1-handle cancellation.
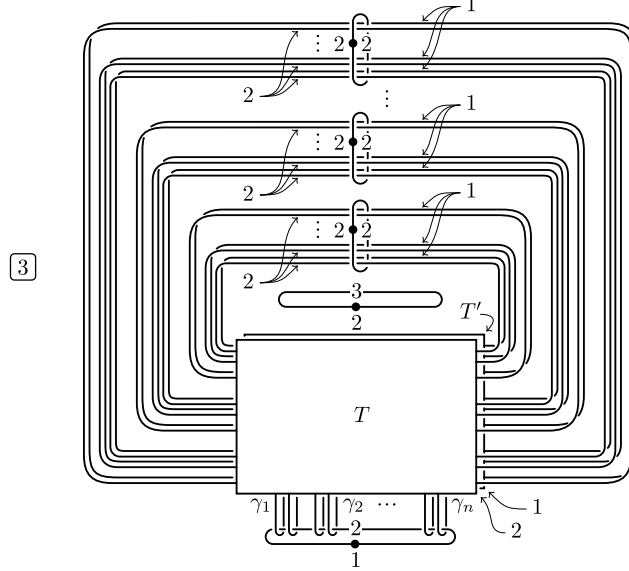

FIGURE 71.

Consider now the framings. To verify that the final framings we obtain coincide with the original ones, we proceed like in the last part of the proof of Proposition 3.1. Let f_i be the framing of L_i in K and c_i be the signed number of the crossings of L_i inverted to get L_i', that is $c_i = (w_i - w_i')/2$ where $w_i = \mathrm{Wr}(L_i)$ and $w_i' = \mathrm{Wr}(L_i')$. We observe that the framing of $L_i \#_{\gamma_i} L_i''$ in the diagram of Figure 71 is the band connected sum of two half integer framings along L_i and L_i'', both of which differ from the blackboard framing by $f_i + 2c_i - 2w_i$ half twists. Hence, for $L_i \#_{\gamma_i} L_i''$ we have $f_i + 2c_i - 2w_i$ full twists added to the blackboard framing. After we have performed the 0/1-handle cancellations to reduce the diagram to an ordinary one, the blackboard framing of $L_i \#_{\gamma_i} L_i''$ can be encoded in the usual way by the integer $w_i + w_i'$. So we are done, since $(f_i + 2c_i - 2w_i) + (w_i + w_i') = f_i$. $\square$

For future reference (see Remark 4.4), we observe that the proof of the above Proposition 3.5 still works if we assume that the link L' used in the construction of F_K is any trivial state of L and not necessarily a vertically trivial state of it. In fact, the vertical triviality of L' has been used only to conclude that L'' is trivial and for that the triviality of L' suffices.

PROPOSITION 3.6. *Any orientable 4-dimensional 2-handlebody H with c connected components is 2-equivalent to a special one having generalized Kirby diagram of the form K_F, for some labelled (orientable) ribbon surface $F \subset B^4$ representing H as a simple branched covering of B^4 of degree $3c$.*

Proof. Up to 1-handle sliding and deletion of cancelling 0/1-handles, we can assume that H has one 0-handle in every component. So, it can be represented by a generalized Kirby diagram K which is the disjoint union of c ordinary Kirby diagrams $K_1, \ldots, K_c$, such that each K_i is separated from all the others and is entirely labelled by i. Disregarding these labels, we construct the labelled ribbon surfaces $F_{K_1}, \ldots, F_{K_c}$. Then, we put $F = F_{K_1} \sqcup \ldots \sqcup F_{K_c}$, after the labels $1, 2, 3$ of each F_{K_i} has been replaced respectively by $3i - 2, 3i - 1, 3i$. The 4-dimensional 2-handlebody represented by F as a $3c$-fold branched covering of B^4 can be proved to be 2-equivalent to H, by applying Proposition 3.5 componentwisely. For the orientability of F, we refer to Remark 2.7. □

REMARK 3.7. Notice that in both the above Propositions 3.5 and 3.6, once a suitable 1-handlebody structure is fixed on F, only 1-handle sliding and addition/deletion of cancelling 0/1-handles are needed to get the wanted 4-dimensional 2-handlebody from K_F, up to handle isotopy.

4. The equivalence theorems

This section completes the proof of the four equivalence theorems stated in the Introduction. The first and main step, is to show that 4-dimensional 2-handlebodies up to 2-deformation are bijectively represented, through the map $F \mapsto K_F$, by simply labelled ribbon surfaces up to labelled 1-isotopy, stabilization and ribbon moves R_1 and R_2, besides labelling conjugation (remember that labelling is actually defined only up to conjugation in Σ_d).

Like in the previous section, we first restrict our attention to the connected case and then come back to the general case with Proposition 4.5. Recall that, for the connected case, we also have the map $K \mapsto F_K$, which associates to each ordinary Kirby diagram K a labelled ribbon surface F_K (defined up to labelled 1-isotopy and ribbon moves) representing its 2-equivalence class as a 3-fold simple branched covering of B^4. In the light of Proposition 3.5, we will be done once we prove that such map is surjective up to labelled 1-isotopy, stabilization and ribbon moves R_1 and R_2. This is the aim of Propositions 4.2 and 4.3.

To begin with, we notice that a d-fold simple branched covering of B^4 represented by a labelled ribbon surface $F \subset B^4$ is connected if and only if the transpositions which appear as labels of any diagram of F generate a transitive subgroup of the symmetric group Σ_d. This is trivially equivalent to say that they generate all Σ_d.

In particular, in this case we can use labelled 1-isotopy move I_2 to expand from F a tongue which, after a suitable sequence of ribbon intersections, is labelled with any given transposition $\tau \in \Sigma_d$ on its tip. Passing all the rest of the diagram through the tip of such a tongue and putting everything back in the original position, has the same effect as conjugating all the labels by τ. Hence, any labelling conjugation can be obtained by a suitable labelled 1-isotopy. This is the reason why labelling conjugation does not appear in the statements of our equivalence theorems concerning connected coverings, while it does in Proposition 4.5.

Before going on, we also introduce the following notion of *special position* for a labelled ribbon surface $F \subset B^4$ representing a (possibly disconnected) simple branched covering of B^4. We say that F is in special position if its diagram is

entirely contained in the projection plane except for a finite number of ribbon half twists and of ribbon intersections and crossings as the ones depicted in Figure 72 with i, j, k and l all distinct and $i < j < k$.
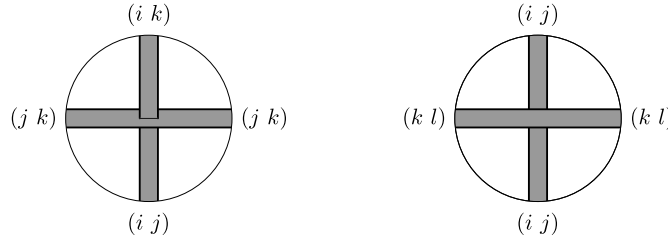
Labelled ribbon surfaces in special position have some remarkable properties that will be useful in the next proofs. Namely, any such F is the disjoint union of subsurfaces $F_1, \ldots, F_{d-1}$, where d is the degree of the covering, such that: 1) the labels attached to F_i are all of the type $(i\ j)$ with $j = i+1, \ldots, d$, and so F_{d-1} is entirely labelled by $(d-1\ d)$; 2) F_i does not form ribbon self intersections or self crossings, that is its diagram can be considered planar except for ribbon half twists; 3) all the ribbon intersections of F consist of a ribbon of F_i which pass through a ribbon of F_j with $i < j$, hence F_1 is nowhere passed through by any other F_i.

Clearly, special position is quite restrictive. For example, even the very peculiar labelled ribbon surfaces F_K are not in special position, due to the ribbon crossings inside the tagle box and to the disks B_i. Nevertheless, the next lemma tells us that things are different if we reason up to ribbon moves.

LEMMA 4.1. *Any labelled ribbon surface representing a connected simple branched covering of B^4 of degree $d \geq 3$ can be put in special position through labelled 1-isotopy and moves R_1 and R_2.*

Proof. Let $F \subset B^4$ be a labelled ribbon surface as in the statement. Forgetting the labelling restrictions of Figure 72, labelled diagram isotopy allows us to make the diagram of F entirely contained in the projection plane except for a finite number of ribbon half twists and of ribbon intersections and crossings. We omit the details of this essentially trivial step and focus on the task of eliminating the ribbon intersections and crossings which do not satisfy the above labelling restrictions.

We change any ribbon intersection between ribbons with disjoint monodromies into a crossing, by a move R_2. Moreover, we change any crossing between ribbons with non-disjoint monodromies into two ribbon intersections, by the first labelled 1-isotopy move of Figure 73, where k may or may not be equal to j. Then, we apply the second labelled 1-isotopy move of Figure 73, where $k \notin \{i, j\}$, to eliminate all the ribbon intersections between ribbons with the same monodromy. Here, we use the hypotheses that the covering is connected and has degree $d \geq 3$, to get the tongue labelled by $(i\ k)$ on its tip. We choose such a tongue to minimize the number of ribbon intersections and crossings, so that none of these is formed with a ribbon having the same monodromy. As above, we replace any crossing with a ribbon having non-disjoint monodromy by two ribbon intersections and any ribbon intersection with a ribbon having disjoint monodromy by a crossing.
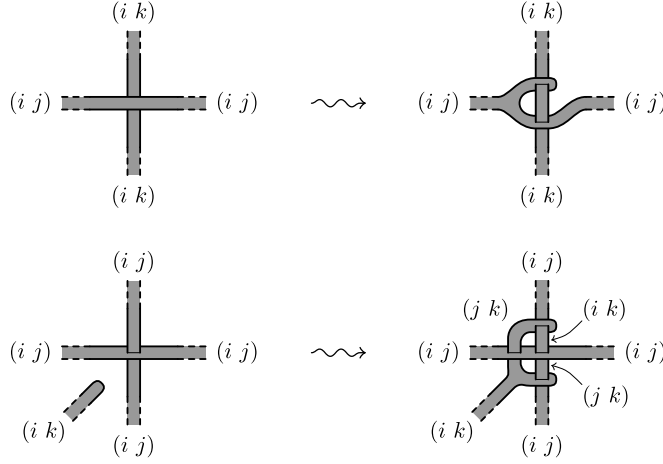
FIGURE 73.

Thus, we are left only with ribbon intersections and crossings as in Figure 72, with i, j, k and l all distinct. The ones which do not satisfy the inequalities $i < j < k$ can be eliminated, by performing one move R_5 followed by two moves $R_1^{\pm 1}$ for the ribbon intersections and just one move R_4 for the crossings. □

Let us now pass on to the announced Propositions 4.2 and 4.3.

PROPOSITION 4.2. *Up to labelled 1-isotopy and moves R_1 and R_2, any labelled ribbon surface $F \subset B^4$ representing a connected simple branched covering of B^4 of degree $d \geq 3$ is equivalent to the d-stabilization of a labelled ribbon surface $F' \subset B^4$ representing a simple 3-fold branched covering of B^4.*

Proof. We proceed by induction on d. For $d = 3$ there is nothing to prove. Given F as in the statement with $d > 3$, we prove that it is equivalent to the d-stabilization of a labelled ribbon surface representing a simple branched covering of B^4 of degree $d - 1$.

To prove the inductive step, we first put F in special position, by applying Lemma 4.1, and modify it in such a way that the label $(1\ d)$ does not appear anymore in its diagram.

Notice that, all the labels $(1\ d)$ of F are attached to the subsurface $F_1 \subset F$, consisting of the pieces of F labelled by $(1\ i)$, with $i = 2, \ldots, d$. As we said after the definition of special position, F_1 does not form ribbon self intersections or self crossings and is nowhere passed through by any other component of F. Moverover, no piece of F_1 labelled by $(1\ d)$ is crossed over by any ribbon.

Consider an adapted 1-handlebody decomposition on F such that crossings and half twists only occur along 1-handles. On the 0-handles of F_1 which are labelled by $(1\ d)$, we operate as in Figure 74, where $1 < i < d$. By choosing the tongue labelled $(i\ d)$ to minimize the number of ribbon intersections and crossings, we can preserve special position (cf. proof of Lemma 4.1).

After that, only some segments of 1-handles delimited by ribbon intersections are still labelled by $(1\ d)$, as sketched on the left side of Figure 75, where $1 < i < j < d$. Here, we have two cases, depending on whether the two delimiting ribbons
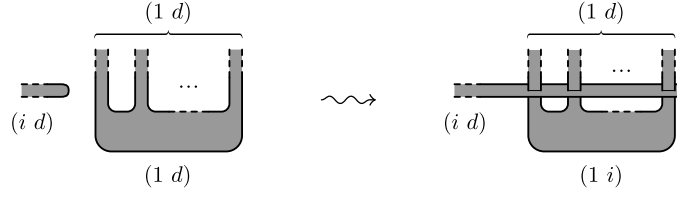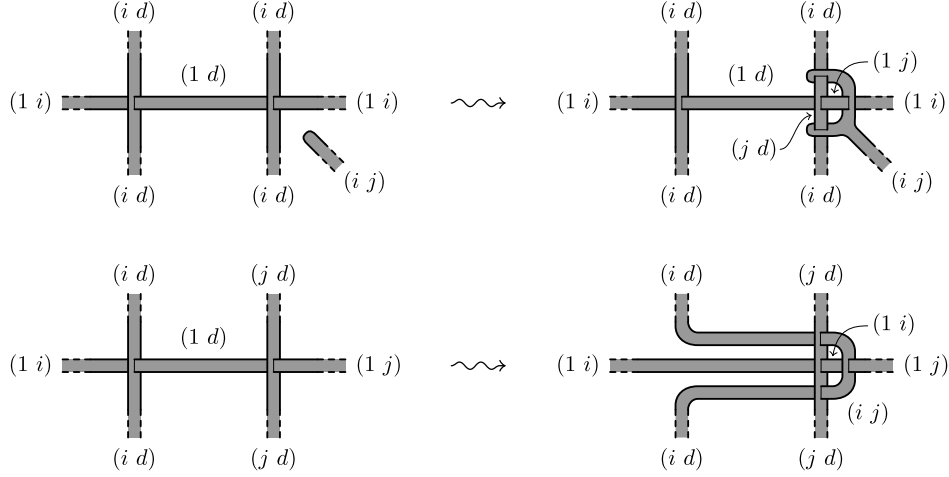
FIGURE 74.

FIGURE 75.

have the same label or not. The upper part of the Figure shows how to reduce the first case to the second, while the lower part tells us how to eliminate the label $(1\ d)$ in this second case. In both cases, we leave to the reader to restore special position and to check that no problem arise with ribbons which possibly cross under the tract labelled by $(1\ d)$.

Once the label $(1\ d)$ has been eliminated from the diagram of F, while preserving special position, we push F_1 down below all the rest of F, except for some tongue terminating at a ribbon intersection, as suggested by right side of Figure 76. This can be done by vertical diagram isotopy and moves R_4 at the ribbon crossings where F_1 crosses above $F - F_1$. Then, we slide F_1 horizontally under $F - F_1$ to make the diagram as in Figure 76, where F_1 is contained in the lower box and $F - F_1$ in the upper one, apart from the ribbons connecting the two boxes. Notice that the labels in the upper (resp. lower) box do not involve 1 (resp. d), while the labels of the connecting ribbons do not involve both 1 and d.
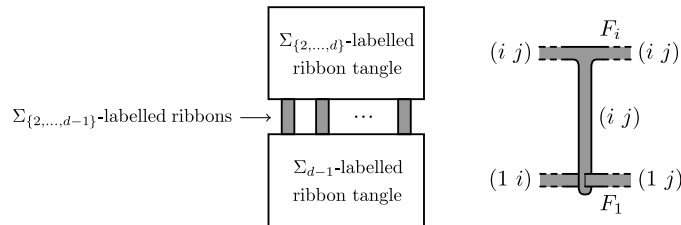

FIGURE 76.

Finally, the modifications described in Figure 77 allows us to isolate a stabilizing disk labelled by $(1\ d)$, by removing d from all the other labels.
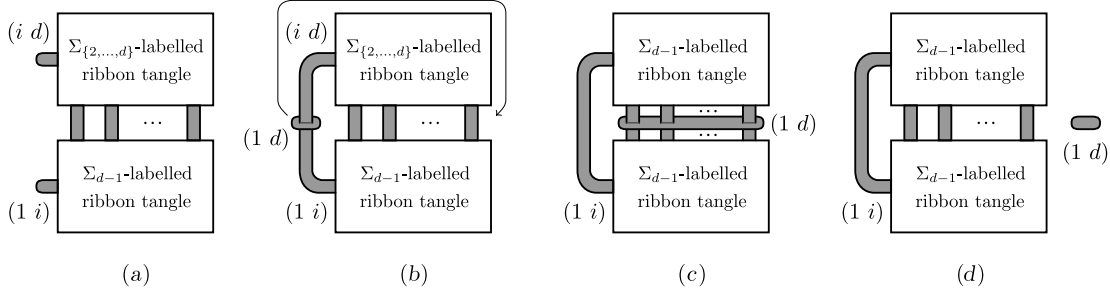

FIGURE 77.

Namely, we expand from the boxes two tongues labelled $(1\ i)$ and $(i\ d)$ for some $i = 2, \ldots, d-1$, as in (a). This can be always done, possibly after having expanded some other $\Sigma_{\{2,\ldots,d-1\}}$-labelled tongues connecting the two boxes, in order to make the traspositions in the upper (resp. lower) box generate all the symmetric group $\Sigma_{\{2,\ldots,d\}}$ (resp. Σ_{d-1}). Then, we connect the tips of the two above tongues by a move R_3 and use labelled 1-isotopy to move the resulting new disk with label $(1\ d)$ as indicated by the arrow in (b). Eventually, we get the diagram in (c), where also the upper box takes labels in Σ_{d-1}, as well as the lower one, so that the only label involving d is the one of the disk between the two boxes. Such disk can be disentagled from the ribbons connecting the boxes by using move R_2, to get (d). $\square$

PROPOSITION 4.3. *For any labelled ribbon surface $F \subset B^4$ representing a connected 3-fold simple branched covering of B^4, there exists an ordinary Kirby diagram K such that the 4-stabilizations of F and F_K are equivalent up to labelled 1-isotopy and moves R_1 and R_2.*

Proof. By Lemma 4.1, we can suppose F to be in special position. In this case, as we said after the definition of special position, F is the disjoint union of two non-empty subsurfaces F_1 and F_2, the first of which takes labels $(1\ 2)$ and $(1\ 3)$, while the second one is entirely labelled by $(2\ 3)$. Moreover, the diagram of F cannot have any ribbon crossing, since there are no disjoint transpositions in Σ_3, and all the ribbon intersections are formed by F_1 passing through F_2. These can be polarized to have planar projection as in the left side of Figure 72 with $i = 1$, $j = 2$ and $k = 3$, up to labelled diagram isotopy which locally half twists the horizontal ribbon.

Consider an adapted 1-handlebody decomposition of F such that half twists only occur along 1-handles. By move I_2 and the tongue technique already seen in the previous proofs, we insert a ribbon intersection along each 1-handle of F_2, taking care that special position is preserved. Then, we apply a move R_5 at every ribbon intersection of F. After that, F_2 is a disjoint union of disks and we can use move R_6 to flatten its diagram into the projection plane, still preserving special position and the above polarization of the ribbon intersections. Finally, a labelled diagram isotopy suffices to put F into the form depicted in Figure 78 (a). Such an isotopy can be realized in two steps: 1) lift all the $(1\ 3)$-labelled parts of F_1 above the projection plane and push all the $(1\ 2)$-labelled ones below it, by a vertical isotopy fixing F_2;
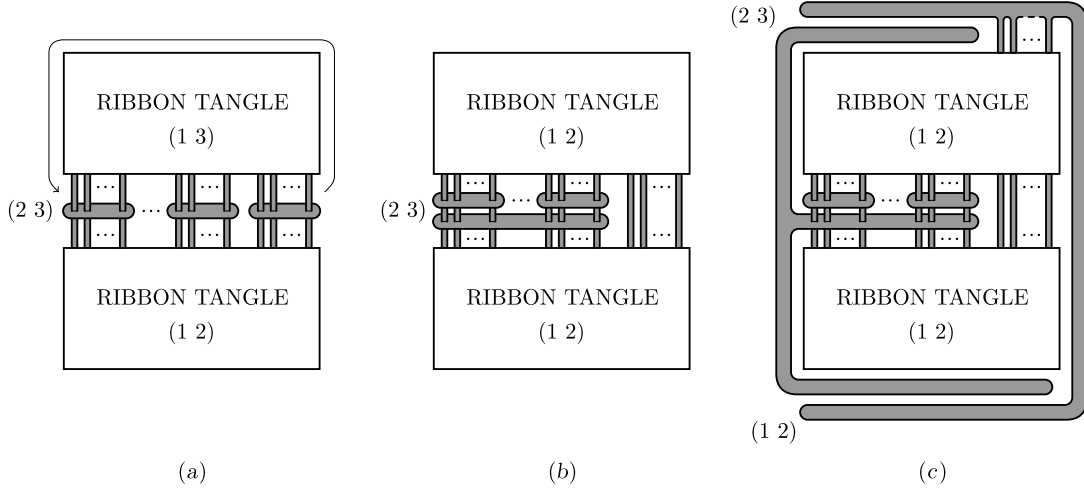
FIGURE 78.

2) move the planar diagram of F to the wanted form, by a suitable horizontal labelled isotopy. Of course, this last step does not preserve any more the special position.

Let us assume that both tangle boxes in Figure 78 (a) are non-empty and that there are at least two (2 3)-labelled disks between them. We leave to the reader to see that such assumption can be made without loss of generality.

By labelled 1-isotopy, we move the rightmost (2 3)-labelled disk as suggested by the arrow, to form a long bar under the other ones like in (b). During this process all the labels in the upper box are changed in (1 2). Then, we obtain the four bars at top and bottom which appear in (c) by labelled diagram isotopy. In particular, the ones labelled by (1 2) are expanded from a 0-handle of F picked up from the upper box.

We warn the reader that the groupings of the vertical bands at different levels in Figure 78 (c), as well as in Figure 80 below, are totally uncorrelated. Their apparent correspondence in the diagrams has only a pictorial value.

The following Figure 79 shows how to incorporate all the 0-handles in the lower box of Figure 78 (c) into the (1 2)-labelled bar at bottom. Here, apart from labelled 1-isotopy, only one move R_3 occurs between (b) and (c). Similarly, all the 0-handles in the upper box can be incorporated into the (1 2)-labelled bar at top. After that, the two ribbon tangles consist of a certain num-
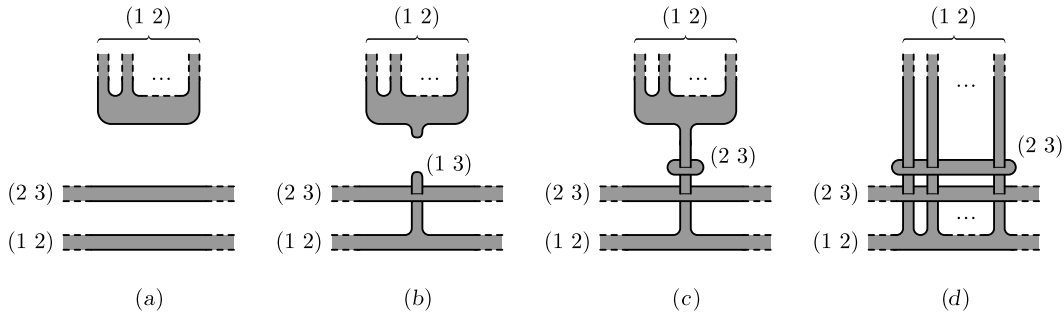

FIGURE 79.

ber of bands which are attached directly to the top/bottom of the (1 2)-labelled bar. We subdivide such bands, by inserting new 0-handles at the intermediate minima and maxima, in such a way that each one of the resulting pieces runs monotonically with respect to the vertical direction of the diagram plane. By labelled diagram isotopy, all the 0-handles corresponding to minima (resp. maxima) inside upper (resp. lower) box can be moved to the lower (resp. upper) one. Then, also the new 0-handles can be incorporated into the bars at top and bottom to get a diagram as in Figure 80 (a), where the ribbon tangles of Figure 78 (c) are replaced by ribbon braids. Denote by X and Y the corresponding ordinary braids, disregarding the ribbon half twists (cf. diagram (b) of Figure 80).
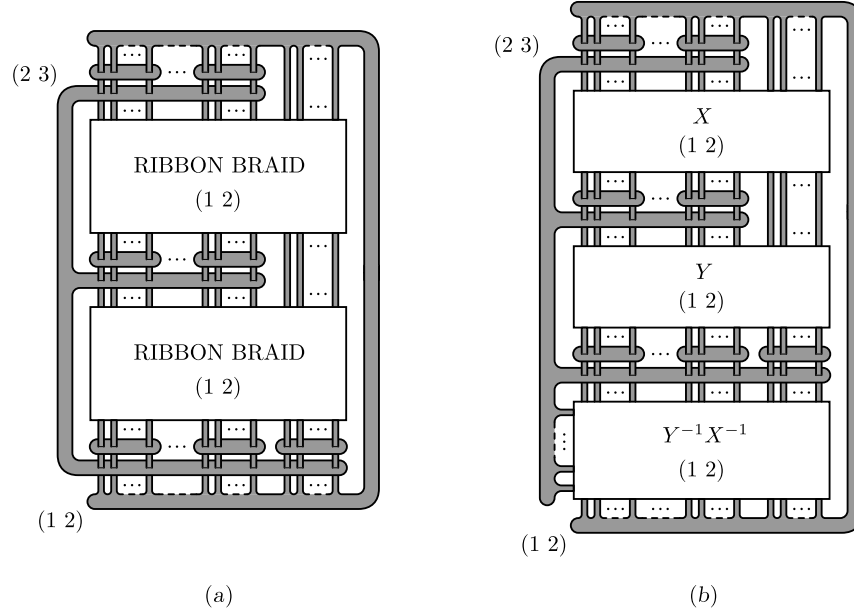
FIGURE 80.

Our next goal is to insert in the diagram a third box with a ribbon braid representing the blackboard framing of $Y^{-1}X^{-1}$, as in Figure 80 (b). The ribbon crossing relative to a standard generator of the braid group can be added just above the bottom bar, together with a small disk expanded from the (2 3)-labelled vertical bar on the left side, as shown in Figure 81. Such a modification essentially coincides with the one described in Figure 65, thus we already know how to realize it in terms of labelled 1-isotopy and ribbon moves. The inverse generator can be dealt with similarly. That is enough to get Figure 80 (b).
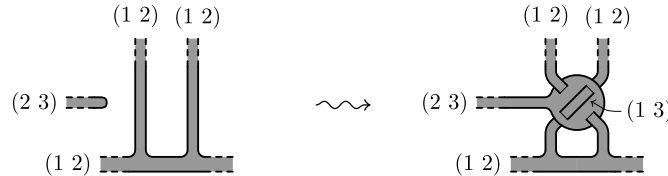
FIGURE 81.

Now, by a labelled diagram isotopy, we align the (2 3)-labelled horizontal bars as sketched in Figure 82. Here, the bars are numbered to make clear the isotopy and, for the sake of readability, only one of the bands forming the ribbon braid of Figure 80 (b) is drawn. During the isotopy, all the other bands are kept parallel to this one outside the boxes. Looking at the right side of Figure 80, we see that the isotopy destroys the braid structure, by introducing self-crossings along the bands. However, such self-crossings are at most six for each band and satisfy a property that will be crucial in the following: going from bottom to top, at each self-crossing the band passes first under and then over.
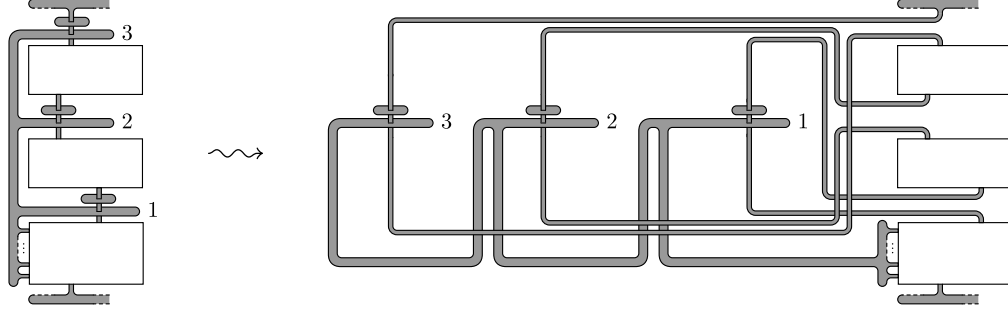
FIGURE 82.

Using the stabilization disk as in Figure 60, the bars we have just aligned can be disconnected in turn from the rest of the diagram and then reconnected differently to form a single long bar like in Figure 83 (a). The tangle box in this last diagram includes the three braid boxes of the previous one together with the obvious part of the bands outside them. To get the subsequent diagram of Figure 83 (b), we simply perform a 90° clockwise rotation and contract the (1 2)-labelled bar (sliding consequently all the bands connecting it to the box).
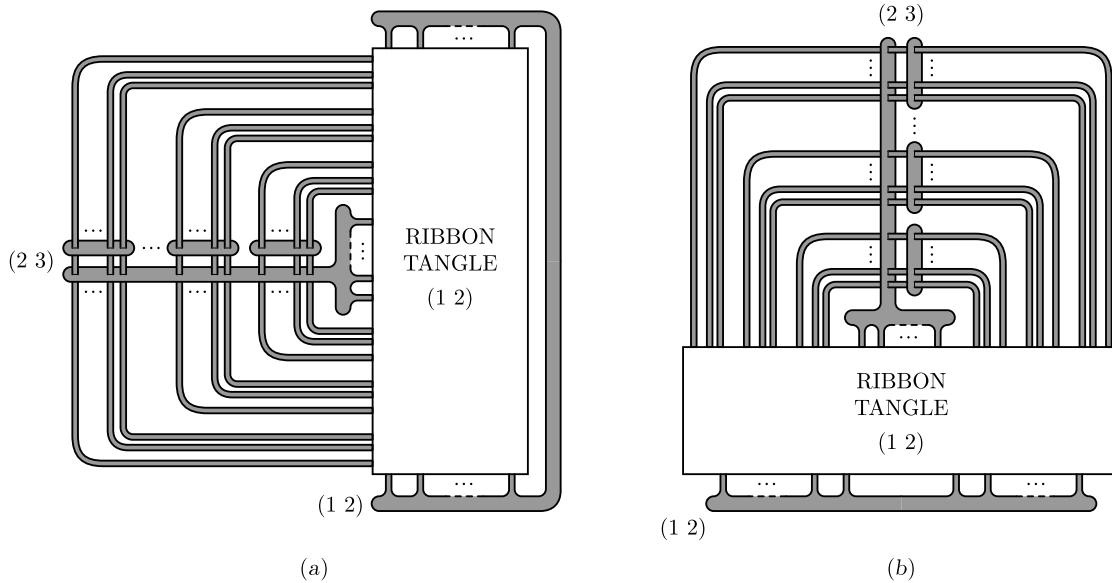
FIGURE 83.
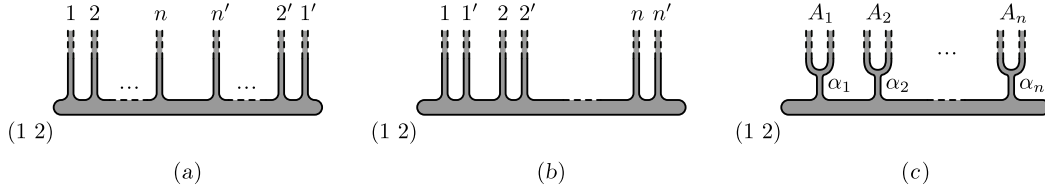
FIGURE 84.

Figure 84 (a) indicates the order in which the bands deriving from the strings of the ribbon braid of Figure 80 (b) are attached to the bottom bar in Figure 83 (b). Here, we numbered by i and i' the two ends of the band corresponding to the i-th string. The same modification depicted in Figure 81 we have already used before, enables us to change this order, by pairing the two ends of the same band as in (b). After that, we can think of the i-th band as a (possibly non-orientable) closed ribbon A_i, connected to the bottom bar by a band α_i, as suggested in (c).

Up to sliding the (1 2)-labelled bands entering the tangle box from the top edge to the lateral ones, the resulting diagram looks like the one of Figure 48, except for the fact that the link $L' = L'_1 \cup \ldots \cup L'_n$ formed by the cores of the A_i's may not be vertically trivial. Indeed, the triviality of the braid in Figure 80 (b) implies that L' is trivial, but not necessarily vertically trivial. On the other hand, each single L_i is vertically trivial, since its only self-crossings are the ones introduced in Figure 83 and they satisfy the property pointed out when introducing the figure. Therefore, we only need to worry about vertically separating different L''_i's.

To this end, let us observe that the triviality of L' is enough to construct disjoint disks $D_1, \ldots, D_n$, with the same properties as in the proof of Proposition 3.1. We can use these disks in turn to vertically separate the L''_i's, just like we did there for proving the independence of F_K from the vertical order of the components of L' (cf. Figure 54 and Figure 55).

To conclude the proof, it remains to verify that the α_i's can be put in the right position as prescribed by the definition of F_K. We leave this trivial task to the reader. $\square$

REMARK 4.4. As a consequence of Proposition 4.3, the link L' in step (c) of the construction of F_K at page 35 can be chosen to be any trivial state (not necessarily a vertically trivial one) of the link L consisting of the framed components of K, without losing the well-definedness of F_K up to 4-stabilization, ribbon moves and 1-isotopy. The reason is that, by Proposition 3.5 and the observation immediately following its proof, the resulting labelled ribbon surface F_K does represent the same 2-deformation class of 4-dimensional 2-handlebodies as if we had chosen L' to be a vertically trivial state of L. Then, Propositions 4.3 and 3.3 allow us to conclude in a straigthforward way.

As an example, in Figure 85 we present the Kirby diagram K and the corresponding ribbon surface F_K for the Akbulut-Kirby 4-sphere Σ_n with $n = 3$. The Kirby diagram is the same as the one drawn in figure 4 of [13], where it is shown that Σ_n is diffeomorphic to B^4 for any n and it is also conjectured that it is not 2-equivalent to B^4 for $n \geq 3$. In the light of Remark 4.4, since in K the link L is already trivial, to obtain F_K we only need to thicken the undotted link components

to (1 2)-labelled ribbons with the right framings, then replace each dotted component with a pair of parallel disks labelled by (2 3), and finally connect by disjoint bands the two (1 2)-labelled components and two of the (2 3)-labelled disks, one for each pair.
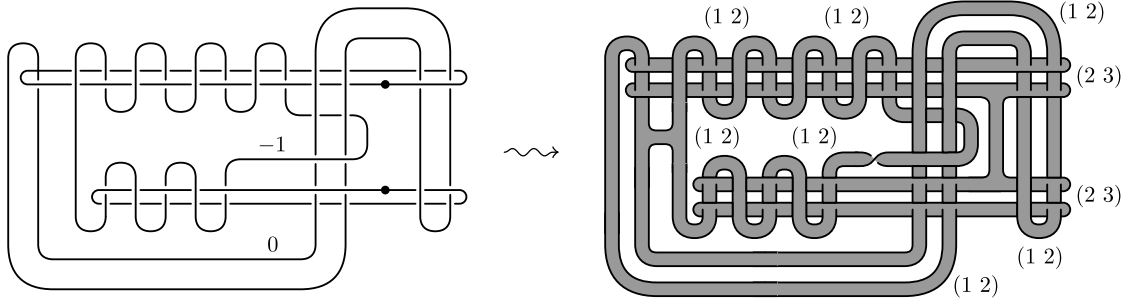

Figure 85.

At this point, we are ready to prove our first equivalence theorem. Actually, this is exclusively a matter of collecting the results we have already got.

Proof of Theorem 1. The "if" part of the theorem is a special case of Propositions 2.4. The "only if" part immediately follows from Propositions 4.2, 4.3, 3.5 and 3.3. Namely, given two labelled ribbon surfaces F and F' representing 2-equivalent 4-dimensional 2-handlebodies as branched covering of B^4 of the same degree $d \geq 4$, we can apply Propositions 4.2 and 4.3 to transform them into d-stabilizations of certain F_K and $F_{K'}$, through moves R_1 and R_2. By the "if" part of the theorem and Proposition 3.5, the two Kirby diagrams K and K' are 2-equivalent. Hence, by Proposition 3.3 F_K and $F_{K'}$ are related by moves R_1 and R_2. □

Before of going on to prove the other equivalence theorems, let us consider the following proposition, which summarizes all we have said until now about branched covering representation of (possibly disconnected) 4-dimensional 2-handlebodies.

PROPOSITION 4.5. *The map $F \mapsto K_F$ induces a bijective correspondence between 4-dimensional 2-handlebodies up to 2-deformation and labelled (orientable) ribbon surfaces, representing them as simple branched coverings of B^4, up to labelling conjugation, labelled 1-isotopy, stabilization and ribbon moves R_1 and R_2. In particular, for handlebodies with c connected components, the coverings can be assumed to have degree $\leq 3c$ and two such coverings representations of 2-equivalent handlebodies can be related involving only coverings of degree $\leq 3c + 1$.*

Proof. By Proposition 3.6, we already know that the correspondence in the statement is surjective, being any 4-dimensional 2-handlebody with c connected components a $3c$-fold simple covering of B^4 branched over a ribbon surface (that can be made orientable by Remark 2.7).

To prove the injectivity, let us consider two coverings representing 2-equivalent 2-handlebodies. Since 2-deformation preserves connectedness, there is a bijective correspondence between the components of the two handlebodies such that corresponding components are 2-equivalent. Up to labelling conjugation, we can assume that the sheets of the two coverings forming the corresponding components are equally numbered. Moreover, by Proposition 4.2 and destabilization, we can reduce

to 3 the maximum number of sheets for each component. Then, we can apply Theorem 1 to each pair of corresponding components in turn, leaving unchanged the other ones. In this way, if the original coverings have degree $\leq 3c$, then all the intermediate coverings involved in relating them have degree $\leq 3c + 1$. $\square$

Having established our main result about branched covering representation of 4-dimensional 2-handlebodies, we pass to prove theorem 2 concerning the case when they have diffeomorphic boundaries.

Proof of Theorem 2. As we observed in the Introduction, moves $P_{\pm}^{\pm 1}$ and $T^{\pm 1}$ do not change the labelled boundary link up to labelled isotopy, so that they also preserve the boundary of the covering manifold up to diffeomorphism. Thus, taking into account Theorem 1 and Proposition 1.4 (b), we only need to show that such moves can be used to interpret blowing up/down and 1/2-handle trading (cf. Figure 19) for an ordinary Kirby diagram K in terms of the labelled ribbon surface F_K. Without loss of generality, we can assume K to be in standard form.

By definition of F_K, it is clear that moves $P_{\pm}$ obviously correspond positive and negative blowups. Figure 86 describes the sequence of moves needed in order to replace the disk C_i corresponding to the i-th 1-handle of K with the ribbon A_{n+1} representing the new 2-handle deriving from the trading.
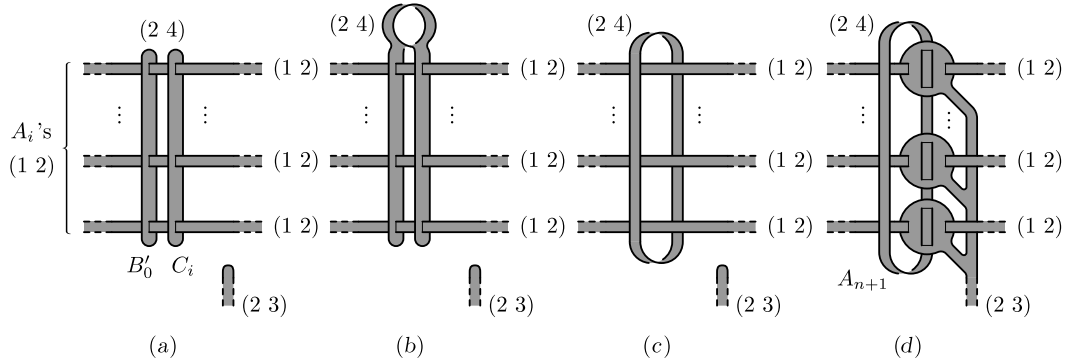
FIGURE 86.

Diagram (a) consists of the disk C_i, after we have operated on it as in Figure 60, together with the parallel disk B_0' resulting from that modification. We perform a move T and labelled 1-isotopy respectively to obtain (b) and (c). This gives us an annulus representing the trivially framed attaching loop of the new 2-handle. Then, we arrange the ribbon surface like in (d), according to a vertically trivial status of the framed link, by inserting some small (2 3)-labelled disks as in 55. Finally, we join the resulting annulus A_{n+1} to A_0, by creating a new band α_{n+1} through a move R_3, and we restore the stabilizing disk, as we did in the proof of Proposition 3.1 by reversing the process of Figure 54. $\square$

Our next goal is to derive Theorem 3 from Theorem 2. The crucial point here is that any simply labelled link in S^3 can be transformed through Montesinos moves into the boundary of a simply labelled ribbon surface in B^4 (see Proposition 4.7). This follows quite directly from Theorem B of [35] about liftable braids, which we state here as Lemma 4.6 after having recalled a couple of definitions.

A simply labelled braid is called a *liftable braid* when the two labellings at its ends coincide. By an *interval* we mean any braid that is conjugate to a standard generator in the braid group. Actually, to make both the terms "liftable" and "interval" meaningful, one should think of braids as self-homeomorphisms of the disk in the usual way (see [6] or [35]), but this is not relevant in the present context.

Of course, a labelled interval, as well as a standard generator, may or may not be liftable depending on the labelling. We say that a labelled interval x is of *type i* if x^i is the first positive power of x which is liftable. It is not difficult to realize that conjugation preserves interval types and that each interval is of type 1, 2 or 3 (cf. Lemma 2.4 of [6] or Lemma 2.3 of [35]).

The labelled intervals x, y and z, whose first liftable positive powers are depicted in Figure 87, are the standard models for the three types above. Namely, any labelled interval of type 1, 2 or 3 is respectively a conjugate of $x^{\pm 1}$, $y^{\pm 1}$ or $z^{\pm 1}$. Evidently, in the figure only the two non-trivial strings of each labelled braid are drawn, the other ones being just horizontal arcs with arbitrary labels. Moreover, in the labelling of each single braid, we assume that i, j, k and l are all different.
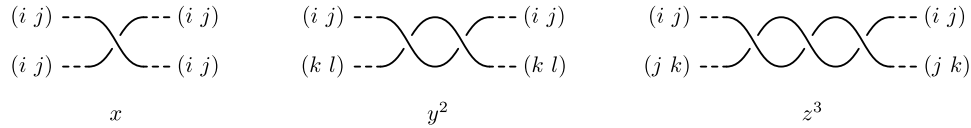

FIGURE 87.

The main result of [35] is the lemma below, which essentially says that any liftable braid is a product of conjugates of labelled braids like the ones in Figure 87.

LEMMA 4.6. *Any liftable braid is a product liftable powers of intervals.*

We emphasize that the lemma holds without restrictions on the degree d of the labelling. However, it is worth observing that the case of $d = 2$ is trivial (every braid is liftable in this case), while the case of $d = 3$ differs from the general one for the absence of intervals of type 2. This special case was previously proved in [6] (cf. also [7]), but the proof of Lemma 4.6 given in [35] does not depend on [6].

The relevant consequence of Lemma 4.6 in the present context is the following branched covering counterpart of the vanishing of the oriented cobordism group Ω_3.

PROPOSITION 4.7. *Any labelled link $L \subset S^3$ representing a (possibly disconnected) d-fold simple branched covering of S^3 is equivalent, up to labelled isotopy and moves M_1 and M_2, to the boundary of labelled ribbon surface $F \subset B^4$ representing a d-fold simple branched covering of B^4.*

Proof. Up to labelled isotopy, we can assume that the link L is the closure $\widehat{B}$ of simply labelled braid B (for example, we can use the labelled version of the well known Alexander's braiding procedure). Of course, B has to be a liftable braid. Then, Lemma 4.6 tells us that, up to labelled isotopy, we can think of B a product of conjugates of braids like $x^{\pm 1}$, $y^{\pm 2}$ or $z^{\pm 3}$ (see Figure 87). Since braids $y^{\pm 2}$ and $z^{\pm 3}$ can be obviously trivialized respectively by moves M_2 and $M_1^{\mp 1}$, we can reduce ourselves to the case when B is a product of liftable intervals.

In this case, a simply labelled ribbon surface $F \subset B^4$ bounded by L can be easily constructed from the band presentation of B (see [40, 41]) determined by its factorization into liftable intervals. Namely, we start with a disjoint union of labelled trivial disks in B^4, spanned by the labelled trivial braid obtained from B by trivializing all the terms $x^{\pm 1}$ appearing in the factorization above. Then, we attach to these disks a labelled twisted band for each such term (see Figure 32 for a simple example, where all the liftable intervals are standard generators).

Notice that the 3-dimensional diagram of the resulting surface may or may not form ribbon intersection, depending on the conjugating braids of the liftable intervals in the factorization of B (cf. [40, 41]). In any case, the labelling consistency when attaching the bands is ensured by the liftability of the intervals. $\square$

Proof of Theorem 3. As we said in the Introduction, it has been known for a long time, since the early work of Montesinos, that moves M_1 and M_2 are covering moves. That is they, as well as labelled isotopy and stabilization, do not change the covering manifold up to diffeomorphism (see Section 1 for a proof of this fact). Therefore, nothing more has to be added about the "if" part of the theorem.

The "only if" part follows from Proposition 4.7 and Theorem 2, taking into account that the restriction of moves R_1 and R_2 to the boundary can be realized by moves M_1 and M_2 (see observation before of Theorem 3 in the Introduction), while moves $P_\pm$ and T preserve the boundary up to labelled isotopy. $\square$

Let us conclude this section with the proof of our last equivalence theorem. This is Theorem 4, which extends the previous Theorem 3 to possibly non-simple coverings of S^3 branched over an embedded graph.

Proof of Theorem 4. We have already observed in Section 1 that moves S_1 and S_2 are covering moves, as they are applications of the coherent monodromies merging principle. Hence, we have only to show that they allow us to transform any labelled graph into a simply labelled link. We proceed in two subsequent steps: 1) we make the labelling simple, by performing moves S_1 on the edges; 2) we make the graph into a link, by performing moves S_2 on the vertices.

Let $G \subset R^3$ be a labelled embedded graph, endowed with a given graph structure without loops (that is every edge has distinct endpoints). We make the labelling simple, by operating on the edges of G one by one. Each time, we assume, up to labelled isotopy, that the edge e under consideration is not involved in any crossing. Denoting by $\sigma \in \Sigma_d$ the label of e, we consider a coherent factorizations $\sigma = \tau_1 \ldots \tau_k$ into transpositions (any minimal factorization of σ is coherent). Then, we split e into k edges $e_1, \ldots, e_k$ with the same endpoints, such that e_i is labelled by τ_i, for each $i = 1, \ldots, k$. To do that, we perform $k-1$ moves S_1, which progressively isolate the traspositions τ_i as labels of new edges. Once all egdes of G have been managed in this way, we are left with a simply labelled graph which we still denote by G.

Now, we operate on the vertices of G one by one, in order to make G into a link. Let v be a vertex of G and $e_1, \ldots, e_h$ be the edges of G having v as an endpoint, numbered according to the counterclockwise order in which they appear around v in the planar diagram of G. Since the total monodromy $\tau_1 \ldots \tau_h$ around v must be trivial, h must be even and the edges around v, can be reodered, up to labelled isotopy, in such a way that $\tau_i = \tau_{h-i+1}$, for every $i = 1, \ldots, h/2$. This immediately follows

from the well known classification of the branched coverings of S^2, if one looks at a small 2-sphere around v transversal to G (cf. [5] or [35]). Then, by $h/2 - 1$ applications of move S_2, we replace the vertex v by $h/2$ non-singular vertices $v_1, \ldots, v_{h/2}$, such that v_i is a common endpoint of e_i and e_{h-i+1}, for each $i = 1, \ldots, h/2$. We leave to the reader to verify that the sequence $\tau_1, \ldots, \tau_{h/2}$ is coherent and that this suffices for the needed moves S_2 to be performable. Obviously, after all the singular vertices of G have been replaced by non-singular ones, we are done. $\square$

5. Final remarks

First of all, we emphasize that the maps $F \mapsto K_F$ and $K \mapsto F_K$, introduced respectively in Sections 2 and 3 (see also Remark 4.4), give an effective way to represent 4-dimensional 2-handlebodies up to 2-deformations as simple coverings of B^4 branched over ribbon surfaces, through generalized Kirby diagrams and Kirby calculus.

Effectiveness is preserved when passing to 3-manifolds too. In particular, being the proof of Lemma 4.6 in [35] constructive, Proposition 4.7 and Theorem 4 (together with the map $F \mapsto K_F$) enable us to define a procedure for obtaining a surgery description of a closed orientable 3-manifold from any presention of it as a branched covering of S^3 (cf. [15] and [16] for the 3-fold case).

Thus, it seems reasonable to expect recognition algorithms and effectively computable invariants for closed orientable 3-manifolds (cf. [27]), based on branched covering representation of them.

Secondly, we point out that our results, other than a different approach to covering moves independent on [37], [38] and [4], also provide the following new line of proof for the Hirsch-Hilden-Montesinos representation theorem: start with the Alexander theorem [3] to represent any closed oriented 3-manifold by a covering of S^3 branched over the 1-skeleton of a 3-simplex; make such covering simple and its branching set into link, as in the proof of Theorem 4; apply Propositions 4.7 and 4.2 in the order, to lower the degree of the covering.

Hopefully, the same ideas could be useful to make some progress in the branched covering representation of smooth closed 4-manifolds. These are known to be 5-fold simple coverings of S^4 branched over non-singular surfaces (see [38] and [20]), but it is an open problem whether the degree can be lowered from 5 to 4. Moreover, any result on covering moves relating diffeomorphic coverings of S^4 is still missing. Theorem 1 together with the results of [30] could give raise to a likely approach to this problem.

Finally, we conclude with some remarks about the relation between the present work and some open problems in the topology of 4-dimensional 2-handlebodies.

A fundamental problem in 4-dimensional topology is the distinction between homeomorphism and diffeomorphism classes. The only known invariants able to detect such difference are the Donaldson and Seiberg-Witten invariants of smooth closed manifolds. These have also been used (cf. [1] and Theorem 8.3.18 in [14]) to distinguish between homeomorphic but non diffeomorphic 4-dimensional 2-handlebodies in the cases when a standard way of closing them is available. In-

variants defined directly on handlebodies and hopefully in purely topological terms are missing and auspicable.

But there is even more delicate question which naturally arises in the topology of 4-dimensional 2-handlebodies and we have already mentioned in Section 1: is there difference between 2-equivalence classes and diffeomorphism classes? In [13] (cf. the example after Remark 4.4) Gompf conjectures that the answer is yes, and offers a list of possible counterexamples. Of course, detecting this phenomena can not rely any more on invariants of smooth manifolds as the Seiberg-Witten ones.

Using the Hennings framework, in [8] have been constructed invariants of 4-dimensional 2-handlebodies under 2-deformations, and these invariants depend on the choice of an unimodular ribbon Hopf algebra.

In the second part of this work [9] we substantially improve this construction, by showing that the map $F \to K_F$ between equivalence classes of ribbon surfaces and Kirby diagrams factors through a bijective map onto the closed morphisms of a universal category $\mathcal{H}^r$. The objects of $\mathcal{H}^r$ form a free $(\otimes, 1)$-algebra on a single object H, and H is required to be a braided ribbon Hopf algebra in $\mathcal{H}^r$. There is a standard procedure of "braiding" a unimodular ribbon Hopf algebra A associating to it a category $\mathcal{H}_A$ and a functor $\mathcal{H}^r \to \mathcal{H}_A$. Therefore the invariants from [8] can be considered as particular examples of the new construction. But the result in [9] is much stronger: it actually gives a complete algebraic description of 4-dimensional 2-handlebodies and we hope is that this would offer new approaches to the open problems in the 4-dimensional topology mentioned above. Moreover, in [9] the result above is used to obtain an analogous algebraic description of the boundaries of 4-dimensional 2-handlebodies, i.e. 3-dimensional manifolds, which resolves for closed manifolds the problem posed by Kerler in [21] (cf. [36, Problem 8-16 (1)]).

Actually, the present work offers yet another possible approach towards studing the difference between 2-deformations and diffeomorphisms: that is by relating it to the difference between 1-isotopy and isotopy of ribbon surfaces.

We recall that 1-isotopy of ribbon surfaces in B^4 was derived from embedded 1-deformation of embedded 2-dimensional 1-handlebodies in B^4, by forgetting the handlebody structure. On the other hand, once one has suitably defined embedded 2-deformation of embedded 2-dimensional 2-handlebodies, isotopy of arbitrary surfaces in B^4 could be derived from it in a similar way. Then, isotopy between ribbon surfaces differs from 1-isotopy just for allowing also addition/deletion of embedded cancelling pairs of 1/2-handles and 2-handle isotopy. This isotopy may involve non-ribbon intersections, such as double loops and triple points, in the diagram.

In different words, we can say that two ribbon surfaces are 1-isotopic if and only if they are isotopic through ribbon surfaces (of course, except for a finite number of intermediate stages whose diagram is not self-tranversal). As we said in Section 1, we do not know whether isotopy relation between ribbon surfaces coincides with 1-isotopy relation or not.

Anologously, since diffeomorphism of 4-dimensional 2-handlebodies is the same as 3-equivalence, the difference between 2-deformations and diffeomorphisms is in the addition/deletion of cancelling pairs of 2/3-handles and 3-handle isotopy. Moreover, the connection established in the previous sections, between labelled 1-isotopy of ribbon surfaces in B^4 and 2-deformation of 4-dimensional 2-handlebodies, through

branched coverings and covering moves, can be at least partially extended. More precisely, attaching a labelled 2-handle to the branching surface $F \subset B^4$ corresponds to attaching a 3-handle to the covering 4-dimensional handlebody H, in such a way that any cancelling pair of 2/3-handles of H can be represented by a cancelling pair of labelled 1/2-handles of F.

A good staring point for studing this problem could be the example of the Akbulut-Kirby sphere Σ_n (see figure 85 for the case of $n = 3$). The proof given in [13] of the fact that Σ_n is diffeomorphic to B^4, is based exactly on the intelligent introduction of a cancelling pair of 2/3-handles, changing the attaching maps by isotopy and eventually cancelling them againts other handles. It would be interesting to see if, at least in this case, these moves correspond to changing the branching surface by isotopy.

References

[1] S. Akbulut, *An exotic 4-manifold*, J. Diff. Geom. **33** (1991), 357–361.

[2] S. Akbulut and R. Kirby, *Branched covers of surfaces in 4-manifolds*, Math. Ann. **252** (1980), 111–131.

[3] J.W. Alexander, *Note on Riemann spaces*, Bull. Amer. Math. Soc. **26** (1920), 370–373.

[4] N. Apostolakis, *On 4-fold covering moves*, Algebraic & Geometric Topology **3** (2003), 117–145.

[5] I. Bernstein and A.L. Edmonds, *On the construction of branched coverings of low-dimensional manifolds*, Trans. Amer. Math. Soc. **247** (1979), 87–124.

[6] J.S. Birman and B. Wajnryb, *3-fold branched coverings and the mappings class group of a surface*, Geometry and Topology, Lecture Notes in Math. **1167**, Springer-Verlag 1985, 24–46.

[7] J. S. Birman and B. Wajnryb, *Presentations of the mapping class group. Errata: "3-fold branched coverings and the mapping class group of a surface" and "A simple presentation of the mapping class group of an orientable surface"*, Israel J. Math. **88** (1994), 425–427.

[8] I. Bobtcheva and M.G. Messia, *HKR-type invariants of 4-thickenings of 2-dimensional CW-complexes*, Alg. and Geom. Top. **3** (2003), 33–87.

[9] I. Bobtcheva and R. Piergallini, *A universal invariant of four-dimensional 2-handlebodies and three-manifolds*, preprint ArXiv:math.GT/0612806.

[10] J. Cerf, *La stratification naturelle des espaces fonction différentiables réelles et la théorème de la pseudo-isotopie*, Publ. Math. I.H.E.S. **39** (1970).

[11] R.H. Fox, *Covering spaces with singularities*, Algebraic Geometry and Topology, A symposium in honour of S. Lefschetz, Princeton 1957, 243–257.

[12] C.A. Giller, *Towards a classical knot theory for surfaces in R^4*, Illinois J. Math. **26** (1982), 591–631.

[13] R.E. Gompf, *Killing the Akbulut-Kirby 4-sphere, with relevance to the Andrews-Curtis and Schoenflies problems*, Topology **30** (1991), 97–115.

[14] R.E. Gompf and A.I. Stipsicz, *4-manifolds and Kirby calculus*, Grad. Studies in Math. **20**, Amer. Math. Soc. 1999.

[15] F. Harou, *Description chirugicale des revêtements triples simples de S^3 ramifiés le long d'un entrelacs*, Ann. Inst. Fourier **51** (2001), 1229–1242.

[16] F. Harou, *Description en terme de revêtements simples de revêtements ramifiés de la sphère*, preprint.

[17] H.M. Hilden, *Every closed orientable 3-manifold is a 3-fold branched covering space of S^3*, Bull. Amer. Math. Soc. **80** (1974), 1243–1244.

[18] H.M. Hilden, *Three-fold branched coverings of S^3*, Bull. Amer. J. Math. **98** (1976), 989–997.

[19] U. Hirsch, *Über offene Abbildungen auf die 3-Sphäre*, Math. Z. **140** (1974), 203–230.

[20] M. Iori and R. Piergallini, *4-manifolds as covers of S^4 branched over non-singular surfaces*, Geometry & Topology **6** (2002), 393–401.

[21] T. Kerler, *Towards an algebraic characterization of 3-dimensional cobordisms*, Contemporary Mathematics **318** (2003), 141–173.

[22] R. Kirby, *A calculus for framed links in S^3*, Invent. math. **45** (1978), 36–56.

[23] R. Kirby, *The topology of 4-manifolds*, Lecture Notes in Mathematics **1374**, Springer-Verlag 1989.

[24] T. Kerler, *Equivalence of bidged links calculus and Kirby's calculus on links on nonsimply connected 3-manifolds*, Topology and its Appl. **87**, (1998), 155–162.

[25] F. Laudenbach and V. Poenaru, *A note on 4-dimensional handlebodies*, Bull. Soc. Math. France **100** (1972), 337–344.

[26] A. Loi and R. Piergallini, *Compact Stein surfaces with boundary as branched covers of S^4*, Invent. math. **143** (2001), 325–348.

[27] S. Matveev, *Algorithmic topology and classification of 3-manifolds*, Algorithms and Computation in Mathematics **9**, Springer 2003.

[28] J.M. Montesinos, *A representation of closed, orientable 3-manifolds as 3-fold branched coverings of S^3*, Bull. Amer. Math. Soc. **80** (1974), 845–846.

[29] J.M. Montesinos, *Three-manifolds as 3-fold branched covers of S^3*, Quart. J. Math. Oxford (2) **27** (1976), 85–94.

[30] J.M. Montesinos, *Heegaard diagrams for closed 4-manifolds*, in "Geometric Topology", J.C. Cantrell ed., Acadecim Press 1979, 219–237.

[31] J.M. Montesinos, *4-manifolds, 3-fold covering spaces and ribbons*, Trans. Amer. Math. Soc. **245** (1978), 453–467.

[32] J.M. Montesinos, *A note on 3-fold branched coverings of S^3*, Math. Proc. Camb. Phil. Soc. **88** (1980), 321–325.

[33] J.M. Montesinos, *Representing 3-manifolds by a universal branching set*, Proc. Camb. Phil. Soc. **94** (1983), 109–123.

[34] J.M. Montesinos, *A note on moves and irregular coverings of S^4*, Contemp. Math. **44** (1985), 345–349.

[35] M. Mulazzani and R. Piergallini, *Lifting braids*, Rend. Ist. Mat. Univ. Trieste **XXXII** (2001), Suppl. 1, 193–219.

[36] T. Ohtsuki, *Problems on invariants of knots and 3-manifolds*, Geom. Topol. Monogr. **4**, in "Invariants of knots and 3-manifolds (Kyoto, 2001)", Geom. Topol. Publ. 2002, 377–572.

[37] R. Piergallini, *Covering Moves*, Trans Amer. Math. Soc. **325** (1991), 903–920.

[38] R. Piergallini, *Four-manifolds as 4-fold branched covers of S^4*, Topology **34** (1995), 497–508.

[39] R. Piergallini and D. Zuddas, *A universal ribbon surface in B^4*, preprint 2003.

[40] L. Rudolph, *Braided surfaces and Seifert ribbons for closed braids*, Comment. Math. Helvetici **58** (1983), 1–37.

[41] L. Rudolph, *Special position for surfaces bounded by closed braids*, Rev. Mat. Ibero-Americana **1** (1985), 93–133; revised version: preprint 2000.

[42] M. Scharlemann, *Unlinking via simultaneous crossing changes*, Trans. Amer. Math. Soc. **336** (1993), 855–868.